

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**

---

**THIS PAGE BLANK (USPTO)**

PF 30-JAN-2001; 2001WO-US00669.  
XX 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-483447/52.  
XX Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human fetal liver -  
XX Claim 27; SEQ ID NO 28569; 639pp + sequence listing; English.  
XX The invention relates to a single exon nucleic acid probe for  
CC measuring human gene expression in a sample derived from human foetal  
CC liver. The single exon nucleic acid probes may be used for predicting,  
CC measuring and displaying gene expression in samples derived from human  
CC fetal liver. The present sequence is a peptide encoded by a single exon  
CC nucleic acid probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX SQ Sequence 74 AA;  
Query Match 100.0%; Score 21; DB 22; Length 74;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SYDA 4  
Db 19 SYDA 22  
RESULT 45  
ABB20450  
ID ABB20450 standard; Protein; 74 AA.  
XX AC ABB20450;  
XX DT 23-JAN-2002 (first entry)  
XX Protein #2449 encoded by probe for measuring heart cell gene expression.  
DE Human; gene expression; heart; microarray; vascular system;  
KW cardiovascular disease; hypertension; cardiac arrhythmia;  
KW congenital heart disease.  
XX OS Homo sapiens.  
XX WO200157274-A2.  
XX PN 09-AUG-2001.  
XX PF 30-JAN-2001; 2001WO-US00666.  
XX 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX PA

XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-488999/53.  
XX Single exon nucleic acid probes for analyzing gene expression in human  
PT hearts -  
XX Claim 15; SEQ ID NO 22220; 530pp; English.  
XX The present invention relates to single exon nucleic acid probes for  
CC measuring human gene expression in a sample derived from human heart (see  
CC ABA21535-APAA1305). The present sequence is a protein encoded by one such  
CC probe. The probes may be used for predicting, measuring and displaying  
CC gene expression in samples derived from the human heart via microarrays.  
CC By measuring gene expression, the probes are useful for predicting,  
CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
CC human heart and vascular system e.g. cardiovascular disease,  
CC hypertension, cardiac arrhythmias and congenital heart disease.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX SQ Sequence 74 AA;  
Query Match 100.0%; Score 21; DB 22; Length 74;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SYDA 4  
Db 19 SYDA 22

Search completed: February 6, 2003, 11:19:38  
Job time : 29.8333 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: February 6, 2003, 11:16:04 : Search time 9.33333 Seconds

(without alignments)  
12.610 Million cell updates/sec

Title: PAT943-3

Sequence: 1 syda 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 106

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCITUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	100.0	7	US-08-310-912A-126	Sequence 126, App
2	21	100.0	7	US-09-301-085-126	Sequence 126, App
3	21	100.0	7	PCT-US95-04589-126	Sequence 126, App
4	21	100.0	14	US-08-557-006C-19	Sequence 19, App
5	21	100.0	22	US-08-630-915A-184	Sequence 184, App
6	21	100.0	25	US-08-930-605-15	Sequence 15, App
7	21	100.0	35	US-08-475-985-3	Sequence 3, App
8	21	100.0	35	US-08-472-172-15	Sequence 15, App
9	21	100.0	35	US-08-256-839-3	Sequence 3, App
10	21	100.0	77	US-09-149-476-441	Sequence 441, App
11	21	100.0	100	US-09-034-916-11	Sequence 11, App
12	21	100.0	132	US-09-325-932A-181	Sequence 181, App
13	21	100.0	133	US-09-325-932A-177	Sequence 177, App
14	21	100.0	204	US-09-615-192A-275	Sequence 275, App
15	21	100.0	237	US-09-325-932A-145	Sequence 145, App
16	21	100.0	243	US-08-578-709-11	Sequence 11, App
17	21	100.0	271	US-09-286-690-2	Sequence 2, App
18	21	100.0	271	US-09-318-794A-2	Sequence 2, App
19	21	100.0	287	US-09-134-001C-5603	Sequence 5603, App
20	21	100.0	290	US-08-936-165A-462	Sequence 462, App
21	21	100.0	316	US-08-038-932B-1	Sequence 1, App
22	21	100.0	316	US-08-656-349-1	Sequence 1, App
23	21	100.0	316	US-08-682-643-4	Sequence 4, App
24	21	100.0	316	US-09-104-623A-5	Sequence 5, App
25	21	100.0	316	US-09-019-532-5	Sequence 5, App
26	21	100.0	316		
27	21	100.0	316		

28	21	100.0	319	US-08-682-643-3	Sequence 3, App
29	21	100.0	331	US-09-134-001C-5254	Sequence 5254, App
30	21	100.0	341	US-09-147-236-2	Sequence 2, App
31	21	100.0	348	US-09-134-001C-4742	Sequence 4742, App
32	21	100.0	376	US-09-134-001C-4259	Sequence 4259, App
33	21	100.0	401	US-09-465-558-70	Sequence 70, App
34	21	100.0	412	US-08-414-938A-4	Sequence 4, App
35	21	100.0	430	US-08-945-848-8	Sequence 8, App
36	21	100.0	432	US-08-472-172-6	Sequence 6, App
37	21	100.0	455	US-08-472-172-4	Sequence 4, App
38	21	100.0	459	US-08-472-172-2	Sequence 2, App
39	21	100.0	487	US-09-004-838-127	Sequence 127, App
40	21	100.0	500	US-08-578-709-15	Sequence 15, App
41	21	100.0	507	US-08-457-274A-22	Sequence 22, App
42	21	100.0	507	PCT-US95-05758-22	Sequence 22, App
43	21	100.0	509	US-08-822-324-8	Sequence 8, App
44	21	100.0	545	US-09-066-046-39	Sequence 39, App
45	21	100.0	552	US-08-557-006C-40	Sequence 40, App

## ALIGNMENTS

RESULT 1  
US-08-310-912A-126  
Sequence 126, Application US/08310912A  
Patent No. 5961730  
GENERAL INFORMATION:  
APPLICANT: Ausubel, Frederick M.  
APPLICANT: Staskawicz, Brian J.  
APPLICANT: Brent, Andrew F.  
APPLICANT: Dahlbeck, Douglas  
APPLICANT: Katagiri, Fumitaki  
APPLICANT: Kunkel, Barbara N.  
APPLICANT: Mandinos, Michael N.  
APPLICANT: Yu, Guo-Liang  
TITLE OF INVENTION: RPS2 GENE FAMILY, PRIMERS, PROBES, AND DETECTION  
TITLE OF INVENTION: METHODS  
NUMBER OF SEQUENCES: 208  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2904  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30B  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/310,912A  
FILING DATE: September 22, 1994  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/227,360  
FILING DATE: April 13, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Lech, Karen F.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/254001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 100254  
INFORMATION FOR SEQ. ID NO: 126:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-310-912A-126

Query Match 100.0%; Score 21; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
||||  
DB 3 SYDA 6

RESULT 2  
US-09-301-085-126  
; Sequence 126, Application US/09301085  
; Patent No. 6262248  
; GENERAL INFORMATION:  
; APPLICANT: Ausubel, Frederick M.  
; APPLICANT: Staskiewicz, Brian J.  
; APPLICANT: Brent, Andrew F.  
; APPLICANT: Dahlbeck, Douglas  
; APPLICANT: Katagiri, Fumiaki  
; APPLICANT: Kunkel, Barbara N.  
; APPLICANT: Mindinos, Michael N.  
; APPLICANT: Yu, Guo-Liang  
; TITLE OF INVENTION: RPS2 GENE FAMILY, PRIMERS, PROBES, AND  
; FILE REFERENCE: 00786/2254002  
; CURRENT APPLICATION NUMBER: US/09/301,085  
; CURRENT FILING DATE: 1999-04-28  
; EARLIER APPLICATION NUMBER: 08/310,912  
; EARLIER FILING DATE: 1994-09-22  
; EARLIER APPLICATION NUMBER: 08/227,360  
; EARLIER FILING DATE: 1994-04-13  
; NUMBER OF SEQ ID NOS: 208  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 126  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-09-301-085-126

Query Match 100.0%; Score 21; DB 4; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
||||  
DB 3 SYDA 6

RESULT 3  
PCT-US95-04589-126  
; Sequence 126, Application PC/TUS9504589  
; GENERAL INFORMATION:  
; APPLICANT: Ausubel, Frederick M.  
; APPLICANT: Staskiewicz, Brian J.  
; APPLICANT: Brent, Andrew F.  
; APPLICANT: Dahlbeck, Douglas  
; APPLICANT: Katagiri, Fumiaki  
; APPLICANT: Kunkel, Barbara N.  
; APPLICANT: Mindinos, Michael N.  
; APPLICANT: Yu, Guo-Liang  
; TITLE OF INVENTION: RPS2 GENE AND USES THEREOF  
; NUMBER OF SEQUENCES: 201  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street Suite 3100  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2904  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30B  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04589  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/227,360  
FILING DATE: 13-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/230001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 100254

INFORMATION FOR SEQ ID NO: 126:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-04589-126

Query Match 100.0%; Score 21; DB 5; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
||||  
DB 3 SYDA 6

RESULT 4  
US-08-557-006C-19  
; Sequence 19, Application US/08557006C  
; Patent No. 6258547  
; GENERAL INFORMATION:  
; APPLICANT: Berli, Rajinder K.  
; APPLICANT: Forster, Robert A.  
; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE  
; FILE REFERENCE: NCAP/PM37588/UST  
; CURRENT APPLICATION NUMBER: US/08/557,006C  
; CURRENT FILING DATE: 1996-03-06  
; PRIOR APPLICATION NUMBER: PCT/GB94/01093  
; PRIOR FILING DATE: 1994-05-20  
; PRIOR APPLICATION NUMBER: GB 9310489.1  
; PRIOR FILING DATE: 1993-05-21  
; PRIOR APPLICATION NUMBER: GB 9318010.7  
; PRIOR FILING DATE: 1993-08-31  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 19  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Rat  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(14)  
US-08-557-006C-19

Query Match 100.0%; Score 21; DB 4; Length 14;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
||||  
DB 5 SYDA 8

RESULT 5

US-08-630-915A-184  
; Sequence 184, Application US/08630915A  
; Patent No. 6309820  
; GENERAL INFORMATION:  
; APPLICANT: SPARKS, Andrew B.  
; APPLICANT: HOFFMAN, No. 6309820h  
; APPLICANT: KAY, Brian K.  
; APPLICANT: FOWLES, Dana M.  
; APPLICANT: MCCONNELL, Stephen J.  
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL  
; TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/630,915A  
; FILING DATE: 03-APR-1996  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mistrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-174  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-8864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 184:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 22 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; US-08-630-915A-184  
  
Query Match 100.0%; Score 21; DB 4; Length 22;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 SYDA 4  
1111  
DB 1 SYDA 4  
  
RESULT 6  
US-08-930-605-15  
; Sequence 15, Application US/08930605  
; Patent No. 5981486  
; GENERAL INFORMATION:  
; APPLICANT: MATSUSHIMA, Kouji  
; APPLICANT: ISHIKAWA, Yuji  
; TITLE OF INVENTION: IKBa PHOSPHORYLATION INHIBITORY PEPTIDE  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLEY & LARDNER  
; STREET: 3000 K Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/930,605  
; FILING DATE: 14-OCT-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-111033  
; FILING DATE: 13-APR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP96/01028  
; FILING DATE: 12-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wegner, Harold C.  
; REGISTRATION NUMBER: 25,258  
; REFERENCE/DOCKET NUMBER: 053466/0215  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 672-5300  
; TELEFAX: (202) 672-5399  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-930-605-15  
  
Query Match 100.0%; Score 21; DB 2; Length 25;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 SYDA 4  
1111  
DB 10 SYDA 13  
  
RESULT 7  
US-08-475-989-3  
; Sequence 3, Application US/08475989  
; Patent No. 5679352  
; GENERAL INFORMATION:  
; APPLICANT: CHONG, Peje  
; APPLICANT: KANDIL, Ali  
; APPLICANT: SIA, Charles  
; APPLICANT: KLEIN, Michel  
; TITLE OF INVENTION: Synthetic Haemophilus Influenzae  
; TITLE OF INVENTION: Conjugate Vaccine  
; NUMBER OF SEQUENCES: 56  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: Suite 701, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/475,989  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/256,839  
; FILING DATE: 03-FEB-1993  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/CA93/00041

FILING DATE: 03-FEB-1993  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9202219.3  
FILING DATE: 03-FEB-1992  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: STEWART, MICHAEL I.  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-505 MIS:V9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-475-989-3

Query Match 100.0%; Score 21; DB 1; Length 35;  
Best Local Similarity 100.0%; Pred. No. 79;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
Db 24 SYDA 27

RESULT 8  
US-08-475-985-3  
Sequence 3, Application US/08475985  
Patent No. 5972349  
GENERAL INFORMATION:  
APPLICANT: CHONG, Pele  
APPLICANT: KANDIL, Ali  
APPLICANT: SIA, Charles  
APPLICANT: KLEIN, Michel  
TITLE OF INVENTION: Synthetic Haemophilus Influenzae  
TITLE OF INVENTION: Conjugate Vaccine  
NUMBER OF SEQUENCES: 56  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/475,985  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/256,839  
FILING DATE: 03-FEB-1993  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/CA93/00041  
FILING DATE: 03-FEB-1993  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9202219.3  
FILING DATE: 03-FEB-1992  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: STEWART, MICHAEL I.  
REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-506 MIS:V9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-475-985-3

Query Match 100.0%; Score 21; DB 2; Length 35;  
Best Local Similarity 100.0%; Pred. No. 79;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
Db 24 SYDA 27

RESULT 9  
US-08-472-172-15  
Sequence 15, Application US/08472172  
Patent No. 5985288  
GENERAL INFORMATION:  
APPLICANT: Munson, Jr., Robert S  
APPLICANT: Grass, Susan  
APPLICANT: Chong, Pele Y Y  
APPLICANT: Fahim, Raafat  
APPLICANT: SIA, Charles D Y  
APPLICANT: McVerry, Patrick  
APPLICANT: Klein, Michel  
TITLE OF INVENTION: Outer Membrane Protein P1 and Peptides  
TITLE OF INVENTION: of Haemophilus Influenzae Type B  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R6  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,172  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/849,411  
FILING DATE: 07-JUL-1992  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I.  
REGISTRATION NUMBER: 24973  
REFERENCE/DOCKET NUMBER: 1038-471  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
TELEX: 063-24567 SIMBAS  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-472-172-15

Query Match 100.0%; Score 21; DB 2; Length 35;

Best Local Similarity 100.0%; Pred. No. 79;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
|||||  
Db 24 SYDA 27

## RESULT 10

US-08-256-839-3  
; Sequence 3, Application US/08256839  
; Patent No. 6018019  
; GENERAL INFORMATION:  
; APPLICANT: CHONG, Pele  
; APPLICANT: KANDIL, Ali  
; APPLICANT: STA, Charles  
; APPLICANT: KLEIN, Michel  
; TITLE OF INVENTION: Synthetic Haemophilus Influenzae  
; TITLE OF INVENTION: Conjugate Vaccine  
; NUMBER OF SEQUENCES: 56  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Slim & McBurney  
; STREET: Suite 701, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/256,839  
; FILING DATE:  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: STEWART, MICHAEL I.  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-373 MTS:jb  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 35 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-256-839-3

Query Match 100.0%; Score 21; DB 3; Length 35;  
Best Local Similarity 100.0%; Pred. No. 79;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 SYDA 4  
|||||  
Db 24 SYDA 27

## RESULT 11

US-09-149-476-441  
; Sequence 441, Application US/09149476  
; Patent No. 6420526  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 186 Human Secreted proteins  
; FILE REFERENCE: P2002P1  
; CURRENT APPLICATION NUMBER: US/09/149,476  
; CURRENT FILING DATE: 1998-09-08  
; EARLIER APPLICATION NUMBER: PCT/US98/04493  
; EARLIER FILING DATE: 1998-03-06  
; EARLIER APPLICATION NUMBER: 60/040,162

; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,333  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/038,621  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,626  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,334  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,336  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,163  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/047,600  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,615  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,597  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,502  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,633  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,583  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,617  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,618  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,503  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,592  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,581  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,584  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,500  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,587  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,492  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,598  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,613  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,582  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,596  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,612  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,632  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,601  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/043,580  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,568  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,314  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,569  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,311  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,671  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,674  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,669  
; EARLIER FILING DATE: 1997-04-11

```

EARLIER APPLICATION NUMBER: 60/043,312
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,313
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,672
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,315
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/056,886
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,877
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,889
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,893
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,630
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,878
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,662
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,872
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,637
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,903
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23

EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Query Match          100.0%  Score 121;  DB 4;  Length 77;
Best Local Similarity 100.0%  Pred. No. 1.8e+02;
Matches 4;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY      1 SYDA 4
      1111
Db      35 SYDA 38

RESULT 12
US-09-034-916-11
; Sequence 11, Application US/09034916
; Patent No. 6046314
; GENERAL INFORMATION:
; APPLICANT: GEBE, JOHN A.
; APPLICANT: STADAK, ANTHONY W.
; APPLICANT: ARUFFO, ALEJANDRO A.
; TITLE OF INVENTION: SPALPHA: A NOVEL SCAVENGER RECEPTOR
; TITLE OF INVENTION: CYSTEINE-RICH DOMAIN-CONTAINING POLYPEPTIDE AND MONOCLONAL
; TITLE OF INVENTION: ANTIBODIES THEREO
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROBINS & ASSOCIATES
; STREET: 90 MIDDLEFIELD ROAD, SUITE 200
; CITY: MENLO PARK
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
```

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/034,916  
FILING DATE: 04-MAR-1998  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/039,956  
FILING DATE: 06-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINS, ROBERTA L.  
REGISTRATION NUMBER: 33,208  
REFERENCE/DOCKET NUMBER: 5998-0020  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 325-7812  
TELEFAX: (650) 325-7823  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 100 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-034-916-11

Query Match 100.0%; Score 21; DB 3; Length 100;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
||||

DB 31 SYDA 34

#### RESULT 13

US-09-325-932A-181  
Sequence 181, Application US/09325932A

Patent No. 6451604

GENERAL INFORMATION:

APPLICANT: Filinn, Barry

APPLICANT: Lasham, Annette

TITLE OF INVENTION: Compositions affecting programmed cell

TITLE OF INVENTION: death and their use in the modification of forestry plant develo

FILE REFERENCE: 1022

CURRENT APPLICATION NUMBER: US/09/325,932A

CURRENT FILING DATE: 1999-06-04

NUMBER OF SEQ ID NOS: 206

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 181

LENGTH: 132

TYPE: PRT

ORGANISM: Pinus radiata

US-09-325-932A-181

Query Match 100.0%; Score 21; DB 4; Length 132;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
||||

DB 126 SYDA 129

#### RESULT 14

US-09-325-932A-177  
Sequence 177, Application US/09325932A

Patent No. 6451604

GENERAL INFORMATION:

APPLICANT: Filinn, Barry

APPLICANT: Lasham, Annette

TITLE OF INVENTION: Compositions affecting programmed cell

TITLE OF INVENTION: death and their use in the modification of forestry plant dev

FILE REFERENCE: 1022

CURRENT APPLICATION NUMBER: US/09/325,932A

CURRENT FILING DATE: 1999-06-04

NUMBER OF SEQ ID NOS: 206

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 177

LENGTH: 133

TYPE: PRT

ORGANISM: Eucalyptus grandis

US-09-325-932A-177

Query Match 100.0%; Score 21; DB 4; Length 133;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
||||

DB 21 SYDA 24

#### RESULT 15

US-09-615-192A-275  
Sequence 275, Application US/09615192A

Patent No. 6410718

GENERAL INFORMATION:

APPLICANT: Bloksberg, Leonard N.

APPLICANT: Havukkala, Ilkka

TITLE OF INVENTION: Materials and Methods for the

TITLE OF INVENTION: modification of Plant Lignin Content

FILE REFERENCE: 11000,1003c4u

CURRENT APPLICATION NUMBER: US/09/615,192A

CURRENT FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 08/975,316

PRIOR FILING DATE: 1997-11-21

PRIOR APPLICATION NUMBER: US 08/713,000

PRIOR FILING DATE: 1996-09-11

PRIOR APPLICATION NUMBER: US 09/169,789

PRIOR FILING DATE: 1998-10-09

NUMBER OF SEQ ID NOS: 405

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 275

LENGTH: 160

TYPE: PRT

ORGANISM: Pinus radiata

US-09-615-192A-275

Query Match 100.0%; Score 21; DB 4; Length 160;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
||||

DB 78 SYDA 81

US-09-325-932A-145  
Sequence 145, Application US/09325932A

Patent No. 6451604

GENERAL INFORMATION:

APPLICANT: Filinn, Barry

APPLICANT: Lasham, Annette

TITLE OF INVENTION: Compositions affecting programmed cell

TITLE OF INVENTION: death and their use in the modification of forestry plant dev

FILE REFERENCE: 1022

CURRENT APPLICATION NUMBER: US/09/325,932A

CURRENT FILING DATE: 1999-06-04

NUMBER OF SEQ ID NOS: 206

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 145

LENGTH: 204

TYPE: PRT

ORGANISM: EucaIyptus grandis  
US-09-325-932A-145

Query Match  
Best Local Similarity 100.0%; Score 21; DB 4; Length 204;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
1111  
DB 21 SYDA 24

RESULT 17  
US-08-578-709-11

Sequence 11, Application US/08578709  
Patent No. 5814509

GENERAL INFORMATION:

APPLICANT: TANABE, Tadashi

TITLE OF INVENTION: PROSTACYCLIN SYNTHASE DERIVED FROM HUMAN

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS

STREET: 2100 Pennsylvania Avenue, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/578,709

FILING DATE: 28-DEC-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP95/00838

FILING DATE: 27-APR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 114316/1994

FILING DATE: 28-APR-1994

ATTORNEY/AGENT INFORMATION:

NAME: Gubinsky, Louis

REGISTRATION NUMBER: 24,835

REFERENCE/DOCKET NUMBER: Q0439

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)293-7060

TELEFAX: (202)293-7860

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 237 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-578-709-11

Query Match

Best Local Similarity 100.0%; Score 21; DB 2; Length 237;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4

1111

DB 82 SYDA 85

RESULT 18

US-09-286-690-2

Sequence 2, Application US/09286690

Patent No. 6103511

GENERAL INFORMATION:

APPLICANT: Li, Xin-Liang

APPLICANT: Ljungdahl, Lars G.

APPLICANT: Chen, Huizhong

TITLE OF INVENTION: Lichenase and Coding Sequences

FILE REFERENCE: 55-96

CURRENT APPLICATION NUMBER: US/09/286,690

CURRENT FILING DATE: 1999-04-05

EARLIER APPLICATION NUMBER: US 60/027,882

EARLIER FILING DATE: 1999-10-04

EARLIER APPLICATION NUMBER: PCT/US97/17811

EARLIER FILING DATE: 1997-10-03

NUMBER OF SEQ ID NOS: 12

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 2

LENGTH: 245

TYPE: PRT

ORGANISM: Orpiniomyces sp. PC-2

US-09-286-690-2

Query Match

Best Local Similarity 100.0%; Score 21; DB 3; Length 245;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4

1111

DB 239 SYDA 242

RESULT 19

US-09-318-794A-2

Sequence 2, Application US/09318794A

Patent No. 6177264

GENERAL INFORMATION:

APPLICANT: DEGUSSA AKTIENGESLSCHAFT

TITLE OF INVENTION: METHOD FOR THE FERMENTATIVE PRODUCTION OF D-PANTOTHENIC

FILE REFERENCE: eggeling

CURRENT APPLICATION NUMBER: US/09/318,794A

CURRENT FILING DATE: 1999-05-26

PRIOR APPLICATION NUMBER: DE 198 55 312.9

PRIOR FILING DATE: 1998-12-01

NUMBER OF SEQ ID NOS: 9

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 2

LENGTH: 271

TYPE: PRT

ORGANISM: Corynebacterium glutamicum

US-09-318-794A-2

Query Match

Best Local Similarity 100.0%; Score 21; DB 4; Length 271;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4

1111

DB 31 SYDA 34

RESULT 20

US-09-318-793A-4

Sequence 4, Application US/09318793A

Patent No. 6184007

GENERAL INFORMATION:

APPLICANT: Dusch, Nicole

APPLICANT: Kalinowski, Jörn

APPLICANT: Publer, Alfred

TITLE OF INVENTION: METHOD FOR THE FERMENTATIVE PRODUCTION OF D-PANTOTHENIC

FILE REFERENCE: 21123/260204

CURRENT APPLICATION NUMBER: US/09/318,793A

CURRENT FILING DATE: 1999-05-26

NUMBER OF SEQ ID NOS: 13

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 4

LENGTH: 271



TYPE: PRT  
ORGANISM: Corynebacterium glutamicum  
US-09-318-793A-4

Query Match 100.0%; Score 21; DB 4; Length 271;  
Best Local Similarity 100.0%; Pred. No. 6.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
1111  
DB 31 SYDA 34

## RESULT 21

US-09-134-001C-5603  
Sequence 5603, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
FILE REFERENCE: CTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 5603  
LENGTH: 287  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-5603

Query Match 100.0%; Score 21; DB 4; Length 287;  
Best Local Similarity 100.0%; Pred. No. 7.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
1111  
DB 264 SYDA 267

## RESULT 22

US-08-936-165A-462  
Sequence 462, Application US/08936165A  
Patent No. 6348582  
GENERAL INFORMATION:  
APPLICANT: Black, Michael  
APPLICANT: Burnham, Martin  
APPLICANT: Hodgson, John  
APPLICANT: Knowles, David  
APPLICANT: Lonetto, Michael  
APPLICANT: Nicholas, Richard  
APPLICANT: Pratt, Julie  
APPLICANT: Reichard, Richard  
APPLICANT: Rosenberg, Martin  
APPLICANT: Ward, Judith  
TITLE OF INVENTION: No. 6348582e1 Prokaryotic Polynucleotides,  
TITLE OF INVENTION: Polypeptides and Their Uses  
NUMBER OF SEQUENCES: 534  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/936,165A

FILING DATE: 24-SEP-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/027,032

FILING DATE: 24-SEP-1996

ATTORNEY/AGENT INFORMATION:  
NAME: Gimmi, Edward R

REGISTRATION NUMBER: 38,891

REFERENCE/DOCKET NUMBER: P50549

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478

TELEFAX: 610-270-5090

TELEX:

INFORMATION FOR SEQ ID NO: 462:

SEQUENCE CHARACTERISTICS:

LENGTH: 290 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: Protein

US-08-936-165A-462  
Query Match 100.0%; Score 21; DB 4; Length 290;  
Best Local Similarity 100.0%; Pred. No. 7.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
1111  
DB 186 SYDA 189

## RESULT 23

US-08-938-932B-1

Sequence 1, Application US/08038932B

Patent No. 5496710

GENERAL INFORMATION:

APPLICANT: Nagao, Hiromasa

APPLICANT: Yoneya, Takashi

APPLICANT: Miyake, Toshio

APPLICANT: Aoyama, Atsuo

APPLICANT: Kai, Ken-ichi

APPLICANT: Kidokoro, Shun-ichi

APPLICANT: Miki, Yoichihiro

APPLICANT: Endo, Kimiko

APPLICANT: Wada, Akiyoshi

TITLE OF INVENTION: NOVEL PROTEASE

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cushman, Darby & Cushman

STREET: 1100 New York Avenue,

CITY: N.W.

STATE: Washington, D.C.

COUNTRY: U.S.A

ZIP: 20005-3918

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/038,932B

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 4-171479

FILING DATE: 08-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 4-237606

FILING DATE: 14-AUG-1992

PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: JP 4-273413
; FILING DATE: 18-SEP-1992
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-638-4203
; TELEX: 248453
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 316 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-038-932B-1

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 316;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
    ||||
Db 65 SYDA 68

RESULT 24
US-08-656-349-1
; Sequence 1, Application US/08656349
; Patent No. 5728544
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: NOVEL PROTEASE II
; NUMBER OF SEQUENCES: 7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/656,349
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 316 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-656-349-1

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 316;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
    ||||
Db 65 SYDA 68

RESULT 25
US-08-682-643-4
; Sequence 4, Application US/08682643
; Patent No. 6103512
; GENERAL INFORMATION:
; APPLICANT: Venema, Gerhardus; Eljstink, Vincentius
; TITLE OF INVENTION: Thermostable Variants of Neutral
; TITLE OF INVENTION: Proteases of Bacillus Stearothermophilus and
; TITLE OF INVENTION: Bacillus Thermoproteolyticus
; NUMBER OF SEQUENCES: 4
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2" 1.44 Mb Floppy diskette
; COMPUTER: IBM 486/DX100 Valuepoint PC
; OPERATING SYSTEM: Microsoft Windows NT Workstation
; OPERATING SYSTEM: Version 4
; SOFTWARE: PatentIn Release #1.0 Version #1.30 (EPO),
```

```

; SOFTWARE: and Windows No. 6103512epad
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/682,643
; FILING DATE: 19-SEP-96
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/NL95/00037
; FILING DATE: 26-JAN-95
; APPLICATION NUMBER: EP 94200182
; FILING DATE: 27-JAN-94
; ATTORNEY/AGENT INFORMATION:
; NAME: Michaelson, Peter L.
; REGISTRATION NUMBER: 30090
; REFERENCE/DOCKET NUMBER: RUC-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)530-6671
; TELEFAX: (908)530-6584
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 316 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-682-643-4

Query Match
Best Local Similarity 100.0%; Score 21; DB 3; Length 316;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
    ||||
Db 65 SYDA 68

RESULT 26
US-09-104-623A-5
; Sequence 5, Application US/09104623A
; Patent No. 6303752
; GENERAL INFORMATION:
; APPLICANT: Olsen, Arne Agerlin
; APPLICANT: Fatum, Tine Muxoll
; APPLICANT: Deussen, Heinz-Josef
; APPLICANT: Roggen, Erwin Ludo
; TITLE OF INVENTION: A Modified Polypeptide
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: No. 63037520 No. 6303752disk of No. 6303752th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/104,623A
; FILING DATE: 25-JUN-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rozek, Carol
; REGISTRATION NUMBER: 36,993
; REFERENCE/DOCKET NUMBER: 5256.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 316 amino acids
```

```
;
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; STRAIN: Bacillus thermoproteolyticus
US-09-104-623A-5

Query Match
Best Local Similarity 100.0%; Score 21; DB 4; Length 316;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4
   ||||
Db 65 SYDA 68

RESULT 27
US-09-019-532-5
; Sequence 5, Application US/09019532B
; Patent No. 6416756
; GENERAL INFORMATION:
; APPLICANT: Olsen, Arne Agerlin
; APPLICANT: Prent, Annette
; TITLE OF INVENTION: A Modified Enzyme for Skin Care
; FILE REFERENCE: 4922,204-US
; CURRENT APPLICATION NUMBER: US/09/019,532B
; CURRENT FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: 0038/97
; EARLIER FILING DATE: 1997-01-10
; EARLIER APPLICATION NUMBER: 0754/97
; EARLIER FILING DATE: 1997-06-25
; EARLIER APPLICATION NUMBER: 60/051,381
; EARLIER FILING DATE: 1997-07-07
; EARLIER APPLICATION NUMBER: PCT/DK98/00015
; EARLIER FILING DATE: 1998-01-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Bacillus thermoproteolyticus
US-09-019-532-5

Query Match
Best Local Similarity 100.0%; Score 21; DB 4; Length 316;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4
   ||||
Db 65 SYDA 68

RESULT 28
US-08-682-643-3
; Sequence 3, Application US/08682643
; Patent No. 6103512
; GENERAL INFORMATION:
; APPLICANT: Venema, Gerhardus; Eljssink, Vincentius
; TITLE OF INVENTION: Thermostable Variants of Neutral
; TITLE OF INVENTION: Proteases of Bacillus stearothermophilus and
; TITLE OF INVENTION: Bacillus thermoproteolyticus
; NUMBER OF SEQUENCES: 4
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2" 1.44 Mb floppy diskette
; OPERATING SYSTEM: IBM 486/DX100 Valuepoint PC
; OPERATING SYSTEM: Microsoft Windows NT Workstation
; SOFTWARE: Patentin Release #1.0 Version #1.30 (EPO),
; SOFTWARE: and Windows No. 6103512epad
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/682,643
; FILING DATE: 19-SEP-96
```

```
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/NL95/00037
; FILING DATE: 26-JAN-95
; APPLICATION NUMBER: EP 94200182
; FILING DATE: 27-JAN-94
; ATTORNEY/AGENT INFORMATION:
; NAME: Michaelson, Peter L.
; REGISTRATION NUMBER: 30090
; REFERENCE/DOCKET NUMBER: RLK-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)530-6671
; TELEFAX: (908)530-6584
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 319 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-682-643-3

Query Match
Best Local Similarity 100.0%; Score 21; DB 3; Length 319;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4
   ||||
Db 68 SYDA 71

RESULT 29
US-09-134-001C-5254
; Sequence 5254, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5254
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5254

Query Match
Best Local Similarity 100.0%; Score 21; DB 4; Length 331;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4
   ||||
Db 265 SYDA 268

RESULT 30
US-09-147-236-2
; Sequence 2, Application US/09147236A
; Patent No. 6316251
; GENERAL INFORMATION:
; APPLICANT: TONOUCHI, Naoto
; APPLICANT: TSUCHIDA, Takayasu
; APPLICANT: YOSHINAGA, Fumihito
; APPLICANT: TAHARA, Naoki
; APPLICANT: HAYASHI, Takahisa
```

```
; TITLE OF INVENTION: NOVEL GENE, GROUP OF GENES, AND NOVEL BETA-GLUCOSIDASE
; FILE REFERENCE: 6537-011-OPCT
; CURRENT APPLICATION NUMBER: US/09/147,236A
; CURRENT FILING DATE: 1999-04-08
; EARLIER APPLICATION NUMBER: PCT/JP97/03633
; EARLIER FILING DATE: 1997-10-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 341
; TYPE: PRN
; ORGANISM: Acetobacter xylinum
; FEATURE:
; OTHER INFORMATION: n at positions 15741 and 15767 may be a, g, t, or
; OTHER INFORMATION: c
US-09-147-236-2

Query Match
Best Local Similarity 100.0%; Score 21; DB 4; Length 341;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
    |||
Db 242 SYDA 245

RESULT 31
US-09-134-001C-4742
; Sequence 4742, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4742
; LENGTH: 348
; TYPE: PRN
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4742

Query Match
Best Local Similarity 100.0%; Score 21; DB 4; Length 348;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
    |||
Db 291 SYDA 294

RESULT 32
US-09-134-001C-4259
; Sequence 4259, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
```

```
; SEQ ID NO 4259
; LENGTH: 376
; TYPE: PRN
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4259

Query Match
Best Local Similarity 100.0%; Score 21; DB 4; Length 376;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
    |||
Db 106 SYDA 109

RESULT 33
US-09-465-558-70
; Sequence 70, Application US/09465558
; Patent No. 6436557
; GENERAL INFORMATION:
; APPLICANT: Morakinyo, Layo O.
; APPLICANT: Orozco Jr, Emil M.
; TITLE OF INVENTION: TETRAHYDROFOLATE METABOLIC ENZYMES
; FILE REFERENCE: B01322 US NA
; CURRENT APPLICATION NUMBER: US/09/465,558
; CURRENT FILING DATE: 1999-12-17
; EARLIER APPLICATION NUMBER: 60/112,734
; EARLIER FILING DATE: 1998-12-18
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 70
; LENGTH: 401
; TYPE: PRN
; ORGANISM: Trifolium aestivum
US-09-465-558-70

Query Match
Best Local Similarity 100.0%; Score 21; DB 4; Length 401;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
    |||
Db 382 SYDA 385

RESULT 34
US-08-414-938A-4
; Sequence 4, Application US/08414938A
; Patent No. 5994627
; GENERAL INFORMATION:
; APPLICANT: LAGUDAH, EVANS SYLVANUS
; TITLE OF INVENTION: GENETIC SEQUENCES CONFERRING
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: COOPER & DUNHAM LLP
; STREET: 1185 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/414,938A
; FILING DATE: March 31, 1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: JOHN P. WHITE
; REGISTRATION NUMBER: 28,678
```

REFERENCE/DOCKET NUMBER: 0687/47527/JPM/GJG  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: + 212 278 0400  
TELEFAX: + 212 391 0525  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 412 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-414-938A-4

Query Match 100.0%; Score 21; DB 2; Length 412;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
1111  
Db 293 SYDA 296

RESULT 35  
US-08-945-848-8  
Sequence 8, Application US/08945848  
Patent No. 5968772  
GENERAL INFORMATION:  
APPLICANT: MATSUSHIRO, AIZO  
TITLE OF INVENTION: PEARL PROTEIN(MACREIN) AND PROCESS FOR  
TITLE OF INVENTION: THE SAME  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive/6300 Sears Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/945.848  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Cawley, Jr., Thomas A.  
REGISTRATION NUMBER: 40,944  
REFERENCE/DOCKET NUMBER: 19036/34324  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 430 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-945-848-8

Query Match 100.0%; Score 21; DB 2; Length 430;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
1111  
Db 71 SYDA 74

RESULT 36  
US-08-472-172-6

Sequence 6, Application US/08472172  
Patent No. 5985288  
GENERAL INFORMATION:  
APPLICANT: Munson, Jr., Robert S  
APPLICANT: Grass, Susan  
APPLICANT: Chong, Pele Y Y  
APPLICANT: Fahim, Raafat  
APPLICANT: Sia, Charles D Y  
APPLICANT: McVerry, Patrick  
APPLICANT: Klein, Michel  
TITLE OF INVENTION: Outer Membrane Protein P1 and Peptides  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R6  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,172  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/849,411  
FILING DATE: 07-JUL-1992  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24973  
REFERENCE/DOCKET NUMBER: 1038-471  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
TELEX: 063-24567 SIMBAS  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 432 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-472-172-6

Query Match 100.0%; Score 21; DB 2; Length 432;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
1111  
Db 144 SYDA 147

RESULT 37  
US-08-472-172-4  
Sequence 4, Application US/08472172  
Patent No. 5985288  
GENERAL INFORMATION:  
APPLICANT: Munson, Jr., Robert S  
APPLICANT: Grass, Susan  
APPLICANT: Chong, Pele Y Y  
APPLICANT: Fahim, Raafat  
APPLICANT: Sia, Charles D Y  
APPLICANT: McVerry, Patrick  
APPLICANT: Klein, Michel  
TITLE OF INVENTION: Outer Membrane Protein P1 and Peptides  
NUMBER OF SEQUENCES: 26  
US-08-472-172-6

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Slim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,172
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/849,411
FILING DATE: 07-JUL-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24973
REFERENCE/DOCKET NUMBER: 1038-471
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
TELEX: 063-24567 SIMBAS
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 455 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-472-172-4

Query Match      100.0%; Score 21; DB 2; Length 455;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SYDA 4      ||||
Db      144 SYDA 147

RESULT 38
US-08-472-172-2
; Sequence 2, Application US/08472172
; Patent No. 5985288
; GENERAL INFORMATION:
; APPLICANT: Munson, Jr., Robert S
; APPLICANT: Grass, Susan
; APPLICANT: Chong, Pele Y Y
; APPLICANT: Fahm, Raafat
; APPLICANT: Sia, Charles D Y
; APPLICANT: McVerity, Patrick
; APPLICANT: Klehn, Michel
; TITLE OF INVENTION: Outer Membrane Protein P1 and Peptides
; TITLE OF INVENTION: of Haemophilus Influenzae type B
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Slim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

```

1 APPLICATION NUMBER: US/08/472.172
2 FILING DATE: 07-JUN-1995
3 CLASSIFICATION: 435
4 PRIOR APPLICATION DATA:
5 APPLICATION NUMBER: US 07/849,411
6 FILING DATE: 07-JUL-1992
7 CLASSIFICATION: 435
8 ATTORNEY/AGENT INFORMATION:
9 NAME: Stewart, Michael I
10 REGISTRATION NUMBER: 24973
11 REFERENCE/DOCKET NUMBER: 1038-471
12 TELECOMMUNICATION INFORMATION:
13 TELEPHONE: (416) 595-1155
14 TELEFAX: (416) 595-1163
15 TELEX: 063-24567 SIMBAS
16 INFORMATION FOR SEQ ID NO: 2:
17 SEQUENCE CHARACTERISTICS:
18 LENGTH: 459 amino acids
19 TYPE: amino acid
20 STRANDEDNESS: single
21 TOPOLOGY: linear
22 US-08-472-172-2
23
24 Query Match 100.0%; Score 21; DB 2; Length 459;
25 Best Local Similarity 100.0%; Pred. NO. 1.2e+03;
26 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
27
28 QY 1 SYDA 4
29 ||||
30 Db 148 SYDA 151
31
32 RESULT 39
33 US-09-004-838-127
34 : Sequence 127, Application US/09004838
35 : Patent No. 6350933
36 :
37 : GENERAL INFORMATION:
38 :
39 : APPLICANT: Micheltmore, Richard W.
40 :
41 : APPLICANT: Shen, Kathy
42 :
43 : APPLICANT: Meyers, Blake
44 :
45 : TITLE OF INVENTION: Procedures and Materials for
46 : TITLE OF INVENTION: Confering Pest Resistance in Plants
47 :
48 : NUMBER OF SEQUENCES: 140
49 :
50 : CORRESPONDENCE ADDRESSES:
51 :
52 : ADDRESSEE: Townsend and Townsend and Crew LLP
53 :
54 : STREET: Two Embarcadero Center, Eighth Floor
55 :
56 : CITY: San Francisco
57 :
58 : STATE: California
59 :
60 : COUNTRY: USA
61 :
62 : ZIP: 94111-3834
63 :
64 : COMPUTER READABLE FORM:
65 :
66 : MEDIUM TYPE: floppy disk
67 :
68 : COMPUTER: IBM PC compatible
69 :
70 : OPERATING SYSTEM: PC-DOS/MS-DOS
71 :
72 : SOFTWARE: PatentIn Release #1.0, Version #1.30
73 :
74 : CURRENT APPLICATION DATA:
75 :
76 : APPLICATION NUMBER: US/09/004,838
77 :
78 : FILING DATE: 09-JAN-1998
79 :
80 : CLASSIFICATION: 800
81 :
82 : PRIOR APPLICATION DATA:
83 :
84 : APPLICATION NUMBER: US 08/781,734
85 :
86 : FILING DATE: 10-JAN-1997
87 :
88 : ATTORNEY/AGENT INFORMATION:
89 :
90 : NAME: Elmhorn, Gregory P.
91 :
92 : REGISTRATION NUMBER: 38,440
93 :
94 : REFERENCE/DOCKET NUMBER: 023070-07881005
95 :
96 : TELECOMMUNICATION INFORMATION:
97 :
98 : TELEPHONE: (415) 576-0200
99 :
100 : TELEFAX: (415) 576-0300
101 :
102 : INFORMATION FOR SEQ ID NO: 127:
103 :
104 : SEQUENCE CHARACTERISTICS:
105 :
106 : LENGTH: 487 amino acids
107 :
108 : TYPE: amino acid
109 :
110 :

```

STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: -  
LOCATION: 1..487  
OTHER INFORMATION: /note= "RGZT deduced sequence"  
US-09-004-838-127

Query Match 100.0%; Score 21; DB 4; Length 487;  
Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
|||  
Db 144 SYDA 147

RESULT 40  
US-08-578-709-15  
Sequence 15, Application US/08578709  
Patent No. 5814509  
GENERAL INFORMATION:  
APPLICANT: TANABE, Tadashi  
TITLE OF INVENTION: PROSTACYCLIN SYNTHASE DERIVED FROM HUMAN  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SUGHRUE, MION, ZINN, MACPHEAK & SEAS  
STREET: 2100 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/578,709  
FILING DATE: 28-DEC-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP95/00838  
FILING DATE: 27-APR-1995  
PRIOR APPLICATION DATA: JP 114316/1994  
FILING DATE: 28-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Gubinsky, Louis  
REGISTRATION NUMBER: 24,835  
REFERENCE/DOCKET NUMBER: Q40439  
TELEPHONE: (202)293-7060  
TELEFAX: (202)293-7860  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 500 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-578-709-15

Query Match 100.0%; Score 21; DB 2; Length 500;  
Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
|||  
Db 82 SYDA 85

RESULT 41

US-08-457-274A-22  
Sequence 22, Application US/08457274A  
Patent No. 5734086  
GENERAL INFORMATION:  
APPLICANT: Scott, Jeffrey G.  
APPLICANT: Tomita, Takashi  
TITLE OF INVENTION: Cytochrome P4501pr Gene and Its Uses  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
STREET: P.O. Box 1051, Clinton Square  
CITY: Rochester  
STATE: New York  
COUNTRY: USA  
ZIP: 14603

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/457,274A  
FILING DATE:  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Goldman, Michael L.  
REGISTRATION NUMBER: 30,727  
REFERENCE/DOCKET NUMBER: 19603/240 (D-1519)  
TELEPHONE: 716-263-1304  
TELEFAX: 716-263-1600  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 507 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Musca domestica  
STRAIN: Rutgers  
DEVELOPMENTAL STAGE: Adult  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: Chromosome 5  
US-08-457-274A-22

Query Match 100.0%; Score 21; DB 1; Length 507;  
Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
|||  
Db 348 SYDA 351

RESULT 42  
PCT-US95-05758-22  
Sequence 22, Application PC/TUS9505758  
GENERAL INFORMATION:  
APPLICANT: Cornell Research Foundation, Inc.  
TITLE OF INVENTION: Cytochrome P4501pr Gene and Its  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
STREET: P.O. Box 1051, Clinton Square  
CITY: Rochester  
STATE: New York  
COUNTRY: USA  
ZIP: 14603  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/05758  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Goldman, Michael L.  
REGISTRATION NUMBER: 30,727  
REFERENCE/DOCKET NUMBER: 19603/241 (D-1519)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 716-263-1304  
TELEFAX: 716-263-1600  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 507 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Musca domestica  
STRAIN: Rutgers  
DEVELOPMENTAL STAGE: Adult  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: Chromosome 5  
PCT-US95-05758-22

Query Match  
Best Local Similarity 100.0%; Score 21; DB 5; Length 507;  
Pred. No. 1.3e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
DB 348 SYDA 351

RESULT 43  
US-08-822-324-8  
Sequence 8, Application US/08822324  
Patent No. 6129917  
GENERAL INFORMATION:  
APPLICANT: Potempa, Jan S.  
APPLICANT: Travis, James  
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS COMPRISING  
TITLE OF INVENTION: PORPHYROMAN GINGIVALIS PROTEINS AND/OR PEPTIDES AND  
NUMBER OF INVENTION: METHODS  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
STREET: 5370 Manhattan Circle, Suite 201  
CITY: Boulder  
STATE: CO  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/822,324  
FILING DATE: 21-MAR-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/013,945  
FILING DATE: 22-MAR-1996  
ATTORNEY/AGENT INFORMATION:

NAME: Feiber, Donna M.  
REGISTRATION NUMBER: 33,878  
REFERENCE/DOCKET NUMBER: 103-95 WO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 488-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 509 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-822-324-8

Query Match  
Best Local Similarity 100.0%; Score 21; DB 3; Length 509;  
Pred. No. 1.3e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
DB 312 SYDA 315

RESULT 44  
US-09-066-046-39  
Sequence 39, Application US/09066046A  
Patent No. 6204252  
GENERAL INFORMATION:  
APPLICANT: MURPHY, Cheryl  
STOREY, James  
BELTZ, Gerald A.  
COUGHLIN, Richard T.  
TITLE OF INVENTION: CHARACTERIZATION OF GRANULOCYTIC  
ERITROCYTES AND METHODS OF USE  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: HALE AND DORR LLP  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: United States  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/066,046A  
FILING DATE: 24-Apr-1998  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Superko, Colleen  
REGISTRATION NUMBER: 39,850  
REFERENCE/DOCKET NUMBER: 106,941.155  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 526-6000  
TELEFAX: (617) 526-5000  
INFORMATION FOR SEQ ID NO: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 545 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 39:  
US-09-066-046-39

Query Match  
Best Local Similarity 100.0%; Score 21; DB 4; Length 545;  
Pred. No. 1.4e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
DB 311



Db 101 SYDA 104

RESULT 45

US-08-557-006C-40  
; Sequence 40, Application US/08557006C  
; Patent No. 6258547

; GENERAL INFORMATION:

; APPLICANT: Beri, Rajinder K.

; APPLICANT: Catling, David

; APPLICANT: Forde, Robert A.

; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE

; FILE REFERENCE: NCAP/PHM37588/UST

; CURRENT APPLICATION NUMBER: US/08/557,006C

; PCT FILING DATE: 1996-03-06

; PCT APPLICATION NUMBER: PCT/GB94/01093

; PCT FILING DATE: 1994-05-20

; PRIOR APPLICATION NUMBER: GB 9310489.1

; PRIOR FILING DATE: 1993-05-21

; PRIOR APPLICATION NUMBER: GB 9318010.7

; PRIOR FILING DATE: 1993-08-31

; NUMBER OF SEQ ID NOS: 44

; SOFTWARE: Patentln Ver. 2.1

; SEQ ID NO 40

; LENGTH: 552

; TYPE: PRT

; ORGANISM: Rat

; FEATURE:

; NAME/KEY: gene

; LOCATION: (1)..(1747)

; OTHER INFORMATION: Full length cDNA sequence fragment of Human AMPK -

; OTHER INFORMATION: fragment begins at nucleotide 24 and ends with

; OTHER INFORMATION: nucleotide 1765

US-08-557-006C-40

Query Match 100.0%; Score 21; DB 4; Length 552;

Best Local Similarity 100.0%; Pred. No. 1.4e+03;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4

Db 282 SYDA 285

Search completed: February 6, 2003, 11:24:06  
Job time : 11.3333 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 11:22:09 ; Search time 5.83333 Seconds  
(without alignments)  
15.202 Million cell updates/sec

Title: PAT943-3  
Perfect score: 21  
Sequence: 1 syda 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 129505 seqs, 22169297 residues

Total number of hits satisfying chosen parameters: 81

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	100.0	7	10	US-09-867-852-126
2	21	100.0	22	10	US-09-879-957-184
3	21	100.0	65	10	US-09-864-761-34093
4	21	100.0	65	10	US-09-864-761-34172
5	21	100.0	65	10	US-09-864-761-39908
6	21	100.0	74	10	US-09-864-761-35748
7	21	100.0	74	10	US-09-864-761-42544
8	21	100.0	74	10	US-09-864-761-47595
9	21	100.0	114	10	US-09-864-761-45709
10	21	100.0	124	9	US-09-796-692-1351
11	21	100.0	140	10	US-09-970-088-11
12	21	100.0	140	10	US-09-970-088-12
13	21	100.0	140	10	US-09-970-088-13
14	21	100.0	140	10	US-09-970-088-14
15	21	100.0	144	10	US-09-881-752A-306
16	21	100.0	167	10	US-09-864-761-39939
17	21	100.0	175	9	US-09-738-626-4419
18	21	100.0	175	9	US-09-738-626-6309
19	21	100.0	219	9	US-09-738-626-5341

20	21	100.0	248	10	US-09-925-299-958	Sequence 958, App
21	21	100.0	256	9	US-09-993-308-2	Sequence 2, Appl
22	21	100.0	267	9	US-09-808-602-23	Sequence 23, Appl
23	21	100.0	267	9	US-09-808-602-25	Sequence 25, Appl
24	21	100.0	267	9	US-09-808-602-101	Sequence 101, App
25	21	100.0	271	9	US-09-738-626-3633	Sequence 3633, App
26	21	100.0	290	10	US-09-939-980-462	Sequence 462, App
27	21	100.0	321	9	US-09-808-602-106	Sequence 106, App
28	21	100.0	332	9	US-09-738-626-4308	Sequence 4308, App
29	21	100.0	332	10	US-09-962-618-2	Sequence 2, Appl
30	21	100.0	339	10	US-09-815-242-10953	Sequence 10953, A
31	21	100.0	343	10	US-09-815-242-10844	Sequence 10844, A
32	21	100.0	345	9	US-09-738-626-4252	Sequence 4252, App
33	21	100.0	352	9	US-09-808-602-105	Sequence 105, App
34	21	100.0	352	10	US-09-888-615-95	Sequence 95, Appl
35	21	100.0	359	10	US-09-799-777-45	Sequence 45, Appl
36	21	100.0	360	9	US-09-738-626-5241	Sequence 5241, App
37	21	100.0	360	9	US-09-815-242-11378	Sequence 11378, A
38	21	100.0	378	9	US-09-738-626-5738	Sequence 5738, App
39	21	100.0	416	9	US-09-971-536-60	Sequence 60, Appl
40	21	100.0	439	10	US-09-815-242-5696	Sequence 5696, App
41	21	100.0	491	10	US-09-820-893-56	Sequence 56, Appl
42	21	100.0	506	10	US-09-815-242-5045	Sequence 5045, App
43	21	100.0	506	10	US-09-815-242-11757	Sequence 11757, A
44	21	100.0	549	10	US-09-925-302-477	Sequence 477, App
45	21	100.0	552	10	US-09-824-735-4	Sequence 4, Appl

#### ALIGNMENTS

RESULT 1  
US-09-867-852-126  
; Sequence 126, Application US/09867852  
; Patent No. US20020147324A1  
; GENERAL INFORMATION:  
; APPLICANT: Ausubel, Frederick M.  
; APPLICANT: Staskawicz, Brian J.  
; APPLICANT: Brent, Andrew F.  
; APPLICANT: Dahlbeck, Douglas  
; APPLICANT: Katagiri, Fumiaki  
; APPLICANT: Kunkel, Barbara N.  
; APPLICANT: Mundry, Michael N.  
; APPLICANT: Yu, Guo-liang  
; TITLE OF INVENTION: RPS2 GENE FAMILY, PRIMERS, PROBES, AND  
; FILE OF INVENTION: DETECTION METHODS  
; FILE REFERENCE: 00786/254002  
; CURRENT APPLICATION NUMBER: US/09/867,852  
; CURRENT FILING DATE: 2001-05-29  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/301,085  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-28  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/310,912  
; PRIOR FILING DATE: EARLIER FILING DATE: 1994-09-22  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/227,360  
; PRIOR FILING DATE: EARLIER FILING DATE: 1994-04-13  
; NUMBER OF SEQ ID NOS: 208  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 126  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-09-867-852-126  
Query Match 100.0%; Score 21; DB 10; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.1e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SYDA 4  
Db 3 SYDA 6  
RESULT 2

US-09-879-957-184  
; Sequence 184, Application US/09879957  
; Patent No. US20020034755A1  
; GENERAL INFORMATION:  
; APPLICANT: SPARKS, Andrew B.  
; HOFFMAN, No. US20020034755A1h  
; KAY, Brian K.  
; FOWLER, Dana M.  
; MCCONNELL, Stephen J.  
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL  
; DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND  
; USING SAME  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/879,957  
; FILING DATE: 13-Jun-2001  
; CLASSIFICATION: <Unknown>  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/630,915  
; FILING DATE: 03-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mirock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-174  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-8864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 184:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 22 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 184:  
US-09-879-957-184  
Query Match 100.0%; Score 21; DB 10; Length 22;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SYDA 4  
| | | |  
| | | |  
Db 1 SYDA 4  
US-09-864-761-34093  
; Sequence 34093, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEOTIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Acomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263,6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 34093  
; LENGTH: 65  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AP000188.1  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.86  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.8  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1  
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.9  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.4  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4  
; OTHER INFORMATION: SWISSPROT HIT: P22415, EVALU 4.00e-14  
; OTHER INFORMATION: EST\_HUMAN HIT: AW937895.1, EVALU 1.00e-13  
US-09-864-761-34093  
Query Match 100.0%; Score 21; DB 10; Length 65;  
Best Local Similarity 100.0%; Pred. No. 78;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SYDA 4  
| | | |  
| | | |  
Db 28 SYDA 31  
US-09-864-761-34172  
; Sequence 34172, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEOTIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Acomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23

APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aemica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 34172  
LENGTH: 65  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AP000044.1  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN BR474, SIGNAL = 2.7  
OTHER INFORMATION: EST HUMAN HIT: AW937895.1, EVALUATE 1.00e-16  
OTHER INFORMATION: SWISSPROT HIT: P22415, EVALUATE 3.00e-17  
US-09-864-761-34172

Query Match 100.0%; Score 21; DB 10; Length 65;  
Best Local Similarity 100.0%; Pred. No. 78;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
|||||  
DB 36 SYDA 39

RESULT 5  
US-09-864-761-39908  
Sequence 39908, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO  
FILE REFERENCE: Aemica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 39908  
LENGTH: 65  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC004812.1  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.4  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5  
OTHER INFORMATION: EST HUMAN HIT: AW503913.1, EVALUATE 1.00e-29  
OTHER INFORMATION: SWISSPROT HIT: P33892, EVALUATE 4.00e-08  
US-09-864-761-39908

Query Match 100.0%; Score 21; DB 10; Length 65;  
Best Local Similarity 100.0%; Pred. No. 78;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
|||||  
Db 22 SYDA 25

RESULT 6  
US-09-864-761-35748  
; Sequence 35748, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 35748  
; LENGTH: 74  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL035705.14  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.1  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.91  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 8.7  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.8  
; OTHER INFORMATION: SWISSPROT HIT: 009137, EVALU8.00e-40  
; OTHER INFORMATION: EST\_HUMAN HIT: BE677760.1, EVALU8.1.00e-29  
US-09-864-761-35748

Query Match 100.0%; Score 121; DB 10; Length 74;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
|||||  
Db 19 SYDA 22

RESULT 7  
US-09-864-761-42544  
; Sequence 42544, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 42544  
; LENGTH: 74  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL035705.22  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7  
; OTHER INFORMATION: SWISSPROT HIT: 009137, EVALU8.00e-40  
; OTHER INFORMATION: EST\_HUMAN HIT: BE677760.1, EVALU8.1.00e-29

US-09-864-761-42544

Query Match 100.0%; Score 21; DB 10; Length 74;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
||||  
DB 19 SYDA 22

## RESULT 8

US-09-864-761-47595  
; Sequence 47595, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; PRIOR FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 47595  
; LENGTH: 74  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AI035705.16  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1  
; OTHER INFORMATION: EXPRESSED IN LONG, SIGNAL = 1.2  
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.97

OTHER INFORMATION: EST HUMAN HIT: BE677760.1, EVALUATE 1.00e-29  
; OTHER INFORMATION: SWISSPROT HIT: Q09137, EVALUATE 8.00e-40  
US-09-864-761-47595

Query Match 100.0%; Score 21; DB 10; Length 74;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
||||  
DB 19 SYDA 22

## RESULT 9

US-09-864-761-45709  
; Sequence 45709, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO  
; FILE REFERENCE: Aecomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; PRIOR FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 45709  
; LENGTH: 114  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC025863.2  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.99

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.87  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.86  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.89  
OTHER INFORMATION: EST HUMAN HIT: BE958611.1, EVALUE 3.00e-40  
OTHER INFORMATION: SWISSPROT HIT: Q50028, EVALUE 8.90e-01  
US-09-864-761-45709

Query Match 100.0%; Score 21; DB 10; Length 114;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
1111  
DB 73 SYDA 76

RESULT 10  
US-09-796-692-1351

Sequence 1351, Application US/09796692  
Publication No. US20020198362A1  
GENERAL INFORMATION:  
APPLICANT: Gaiger, Alexander  
APPLICANT: Mannion, Jane  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
FILE REFERENCE: 2077.001200  
CURRENT APPLICATION NUMBER: US/09/796,692  
PRIOR FILING DATE: 2001-03-01  
PRIOR APPLICATION NUMBER: 60/186,126  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: 60/190,479  
PRIOR FILING DATE: 2000-03-17  
PRIOR APPLICATION NUMBER: 60/200,545  
PRIOR FILING DATE: 2000-04-27  
PRIOR APPLICATION NUMBER: 60/200,303  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: 60/200,779  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: 60/200,999  
PRIOR FILING DATE: 2000-05-01  
PRIOR APPLICATION NUMBER: 60/202,084  
PRIOR FILING DATE: 2000-05-04  
PRIOR APPLICATION NUMBER: 60/206,201  
PRIOR FILING DATE: 2000-05-22  
PRIOR APPLICATION NUMBER: 60/218,950  
PRIOR FILING DATE: 2000-07-14  
PRIOR APPLICATION NUMBER: 60/222,903  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: 60/223,416  
PRIOR FILING DATE: 2000-08-04  
PRIOR APPLICATION NUMBER: 60/223,378  
PRIOR FILING DATE: 2000-08-07  
NUMBER OF SEQ ID NOS: 9597  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1351  
LENGTH: 124  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-796-692-1351

Query Match 100.0%; Score 21; DB 9; Length 124;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
1111  
DB 107 SYDA 110

RESULT 11  
US-09-970-088-11

Sequence 11, Application US/09970088  
Patent No. US20020151489A1  
GENERAL INFORMATION:  
APPLICANT: GRAVEREAUX, EDWIN C.  
APPLICANT: SILVER, MARCY  
APPLICANT: ISNER, JEFFREY M.  
APPLICANT: YOON, YOUNG-SUP  
TITLE OF INVENTION: USE OF LYMPHANGIOGENIC AGENTS TO TREAT LYMPHATIC  
FILE REFERENCE: 71417/55062  
CURRENT APPLICATION NUMBER: US/09/970,088  
CURRENT FILING DATE: 2001-10-02  
PRIOR APPLICATION NUMBER: 60/237,171  
PRIOR FILING DATE: 2000-10-02  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 11  
LENGTH: 140  
TYPE: PRT  
ORGANISM: Oryctolagus cuniculus  
US-09-970-088-11

Query Match 100.0%; Score 21; DB 10; Length 140;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
1111  
DB 135 SYDA 138

RESULT 12  
US-09-970-088-12  
Sequence 12, Application US/09970088  
Patent No. US20020151489A1  
GENERAL INFORMATION:  
APPLICANT: GRAVEREAUX, EDWIN C.  
APPLICANT: SILVER, MARCY  
APPLICANT: ISNER, JEFFREY M.  
APPLICANT: YOON, YOUNG-SUP  
TITLE OF INVENTION: USE OF LYMPHANGIOGENIC AGENTS TO TREAT LYMPHATIC  
FILE REFERENCE: 71417/55062  
CURRENT APPLICATION NUMBER: US/09/970,088  
CURRENT FILING DATE: 2001-10-02  
PRIOR APPLICATION NUMBER: 60/237,171  
PRIOR FILING DATE: 2000-10-02  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 12  
LENGTH: 140  
TYPE: PRT  
ORGANISM: Bos sp.  
US-09-970-088-12

Query Match 100.0%; Score 21; DB 10; Length 140;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
1111  
DB 135 SYDA 138

RESULT 13  
US-09-970-088-13  
Sequence 13, Application US/09970088  
Patent No. US20020151489A1  
GENERAL INFORMATION:  
APPLICANT: GRAVEREAUX, EDWIN C.  
APPLICANT: SILVER, MARCY  
APPLICANT: ISNER, JEFFREY M.  
APPLICANT: YOON, YOUNG-SUP



```
; TITLE OF INVENTION: USE OF LYMPHANGIOGENIC AGENTS TO TREAT LYMPHATIC
; FILE REFERENCE: 71417/55062
; CURRENT APPLICATION NUMBER: US/09/970,088
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/237,171
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 13
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-088-13

Query Match          100.0%; Score 21; DB 10; Length 140;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
   ||||
Db 135 SYDA 138

RESULT 14
US-09-970-088-14
; Sequence 14, Application US/09970088
; Patent No. US20020151489A1
; GENERAL INFORMATION:
; APPLICANT: GRAVEREAUX, EDWIN C.
; APPLICANT: SILVER, MARCY
; APPLICANT: ISNER, JEFFREY M.
; APPLICANT: YOON, YOUNG-SUP
; TITLE OF INVENTION: USE OF LYMPHANGIOGENIC AGENTS TO TREAT LYMPHATIC
; TITLE OF INVENTION: DISORDERS
; FILE REFERENCE: 71417/55062
; CURRENT APPLICATION NUMBER: US/09/970,088
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/237,171
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 14
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-970-088-14

Query Match          100.0%; Score 21; DB 10; Length 140;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
   ||||
Db 135 SYDA 138

RESULT 15
US-09-881-752A-306
; Sequence 306, Application US/09881752A
; Patent No. US20020115078A1
; GENERAL INFORMATION:
; APPLICANT: Kleantous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20020115078A1 Helicobacter Polypeptides in the
; FILE REFERENCE: 06132/041002
; CURRENT APPLICATION NUMBER: US/09/881,752A
; CURRENT FILING DATE: 2001-06-15
```

```
; PRIOR APPLICATION NUMBER: US 08/833,457
; PRIOR FILING DATE: 1997-04-01
; NUMBER OF SEQ ID NOS: 370
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 306
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-881-752A-306

Query Match          100.0%; Score 21; DB 10; Length 144;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
   ||||
Db 89 SYDA 92

RESULT 16
US-09-864-761-39939
; Sequence 39939, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David R.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO: 39939
; LENGTH: 167
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005841.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
; OTHER INFORMATION: SWISSPROT HIT: Q08050, EVALU6 6.00e-83
; OTHER INFORMATION: EST_HUMAN HIT: AW239326.1, EVALU6 1.00e-76
US-09-738-626-4419
```

```
Query Match
Best Local Similarity 100.0%; Score 21; DB 10; Length 167;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYDA 4
||||
DB 128 SYDA 131
```

```
RESULT 17
US-09-738-626-4419
; Sequence 4419, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4419
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4419
```

```
Query Match
Best Local Similarity 100.0%; Score 21; DB 9; Length 175;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYDA 4
||||
DB 119 SYDA 122
```

```
RESULT 18
US-09-738-626-6309
; Sequence 6309, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
```

```
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6309
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6309
```

```
Query Match
Best Local Similarity 100.0%; Score 21; DB 9; Length 175;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYDA 4
||||
DB 64 SYDA 67
```

```
RESULT 19
US-09-738-626-5341
; Sequence 5341, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5341
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5341
```

```
Query Match
Best Local Similarity 100.0%; Score 21; DB 9; Length 219;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 SYDA 4  
||||  
Db 109 SYDA 112

RESULT 20  
US-09-925-299-958

; Sequence 958, Application US/09925299  
; Patent No. US20020055627A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

; FILE REFERENCE: PA102

; CURRENT APPLICATION NUMBER: US/09/925,299

; PRIOR FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: PCT/US00/05883

; PRIOR FILING DATE: 2000-03-08

; PRIOR APPLICATION NUMBER: 60/124,270

; PRIOR FILING DATE: 1999-03-12

; NUMBER OF SEQ ID NOS: 1556

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 958

; LENGTH: 248

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (7)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-925-299-958

Query Match 100.0%; Score 21; DB 10; Length 248;

Best Local Similarity 100.0%; Pred. No. 2.8e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
||||  
Db 128 SYDA 131

RESULT 21

US-09-993-308-2

; Sequence 2, Application US/09993308

; Patent No. US20020159435A1

; GENERAL INFORMATION:

; APPLICANT: Gordon-Kamm, William J.

; APPLICANT: Lowe, Keith S.

; APPLICANT: Larkins, Brian A.

; APPLICANT: Dilkes, Brian R.

; APPLICANT: Sun, Yuejin

; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides and Uses thereof

; FILE REFERENCE: 1146

; CURRENT APPLICATION NUMBER: US/09/993,308

; PRIOR FILING DATE: 2001-11-06

; PRIOR APPLICATION NUMBER: 60/246,349

; PRIOR FILING DATE: 2000-11-07

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 256

; TYPE: PRT

; ORGANISM: Zea mays

US-09-993-308-2

Query Match 100.0%; Score 21; DB 9; Length 256;

Best Local Similarity 100.0%; Pred. No. 2.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
||||  
Db 138 SYDA 141

RESULT 22  
US-09-808-602-23

; Sequence 23, Application US/09808602  
; Patent No. US20020155115A1

; GENERAL INFORMATION:

; APPLICANT: Vernet, Corine A

; APPLICANT: Fernandes, Elma

; APPLICANT: Shimkets, Richard A

; APPLICANT: Herrman, John L

; APPLICANT: Majumder, Kunud

; APPLICANT: Mishra, Vishnu

; APPLICANT: Mezes, Peter S

; APPLICANT: MacDougall, John

; TITLE OF INVENTION: No. US20020155115A1e1 Proteins and Nucleic Acids Encoding Same

; FILE REFERENCE: 15966-697 CIP

; CURRENT APPLICATION NUMBER: US/09/808,602

; PRIOR FILING DATE: 2001-03-14

; PRIOR APPLICATION NUMBER: 09/800,198

; PRIOR FILING DATE: 2001-03-05

; PRIOR APPLICATION NUMBER: 60/186,596

; PRIOR FILING DATE: 2000-03-03

; NUMBER OF SEQ ID NOS: 114

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 23

; LENGTH: 267

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-808-602-23

Query Match 100.0%; Score 21; DB 9; Length 267;

Best Local Similarity 100.0%; Pred. No. 3.1e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
||||  
Db 157 SYDA 160

RESULT 23

US-09-808-602-25

; Sequence 25, Application US/09808602

; Patent No. US20020155115A1

; GENERAL INFORMATION:

; APPLICANT: Vernet, Corine A

; APPLICANT: Fernandes, Elma

; APPLICANT: Shimkets, Richard A

; APPLICANT: Herrman, John L

; APPLICANT: Majumder, Kunud

; APPLICANT: Mishra, Vishnu

; APPLICANT: Mezes, Peter S

; APPLICANT: MacDougall, John

; TITLE OF INVENTION: No. US20020155115A1e1 Proteins and Nucleic Acids Encoding Same

; FILE REFERENCE: 15966-697 CIP

; CURRENT APPLICATION NUMBER: US/09/808,602

; PRIOR FILING DATE: 2001-03-14

; PRIOR APPLICATION NUMBER: 09/800,198

; PRIOR FILING DATE: 2001-03-05

; PRIOR APPLICATION NUMBER: 60/186,596

; PRIOR FILING DATE: 2000-03-03

; NUMBER OF SEQ ID NOS: 114

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 25

; LENGTH: 267

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-808-602-25

Query Match 100.0%; Score 21; DB 9; Length 267;

Best Local Similarity 100.0%; Pred. No. 3.1e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
||||

Db 157 SYDA 160

## RESULT 24

Sequence 101, Application US/09808602  
Patent No. US2002015115A1  
GENERAL INFORMATION:  
APPLICANT: Vernet, Corine A  
APPLICANT: Fernandes, Elma  
APPLICANT: Shimkets, Richard A  
APPLICANT: Heirman, John L  
APPLICANT: Majumder, Kunud  
APPLICANT: Mishra, Vishnu  
APPLICANT: Mezes, Peter S  
TITLE OF INVENTION: No. US2002015115A1e1 Proteins and Nucleic Acids Encoding Same  
FILE REFERENCE: 15966-697 CIP  
CURRENT APPLICATION NUMBER: US/09/808,602  
CURRENT FILING DATE: 2001-03-14  
PRIOR APPLICATION NUMBER: 09/800,198  
PRIOR FILING DATE: 2001-03-05  
PRIOR APPLICATION NUMBER: 60/186,596  
PRIOR FILING DATE: 2000-03-03  
NUMBER OF SEQ ID NOS: 114  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 101  
LENGTH: 267  
TYPE: PRT  
ORGANISM: Macaca mulatta  
US-09-808-602-101

Query Match 100.0%; Score 21; DB 9; Length 267;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
Db 157 SYDA 160

## RESULT 25

US-09-738-626-3633  
Sequence 3633, Application US/09738626  
Publication No. US20020197605A1  
GENERAL INFORMATION:  
APPLICANT: NAKAGAWA, SATOSHI  
APPLICANT: MIZOGUCHI, HIROSHI  
APPLICANT: ANDO, SEIKO  
APPLICANT: HAYASHI, MIKIRO  
APPLICANT: OCHIAI, KEIKO  
APPLICANT: YOKOI, HARUHIKO  
APPLICANT: TATEISHI, NAOKO  
APPLICANT: SENOH, AKIHIRO  
APPLICANT: IKEDA, MASATO  
APPLICANT: OZAKI, AKIO  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-125  
CURRENT APPLICATION NUMBER: US/09/738,626  
CURRENT FILING DATE: 2000-12-18  
PRIOR APPLICATION NUMBER: JP 99/377484  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: JP 00/159162  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: JP 00/280988  
PRIOR FILING DATE: 2000-08-03  
NUMBER OF SEQ ID NOS: 7059  
SOFTWARE: Patentln Ver. 3.0  
SEQ ID NO 3633  
LENGTH: 271  
TYPE: PRT  
ORGANISM: Corynebacterium glutamicum  
US-09-738-626-3633

Query Match 100.0%; Score 21; DB 9; Length 271;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
Db 31 SYDA 34

## RESULT 26

US-09-939-980-462  
Sequence 462, Application US/09939980  
Patent No. US20020082234A1  
GENERAL INFORMATION:  
APPLICANT: Black, Michael  
APPLICANT: Burnham, Martin  
APPLICANT: Hodgson, John  
APPLICANT: Knowles, David  
APPLICANT: Lonetto, Michael  
APPLICANT: Nicholas, Richard  
APPLICANT: Pratt, Julie  
APPLICANT: Reichard, Richard  
APPLICANT: Rosenberg, Martin  
APPLICANT: Ward, Judith  
TITLE OF INVENTION: No. US20020082234A1e1 Prokaryotic Polynucleotides,  
NUMBER OF SEQUENCES: 534  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Smithline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/939,980  
FILING DATE: 27-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/936,165  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimmil, Edward R  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P50549  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 462:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 290 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 462:  
US-09-939-980-462

Query Match 100.0%; Score 21; DB 10; Length 290;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
Db 186 SYDA 189

## RESULT 27

US-09-808-602-106  
; Sequence 106, Application US/09808602  
; Patent No. US2002015115A1  
; GENERAL INFORMATION:  
; APPLICANT: Vernet, Corine A  
; APPLICANT: Fernandes, Elma  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Herrman, John L  
; APPLICANT: Majumder, Kumud  
; APPLICANT: Mishra, Vishnu  
; APPLICANT: Mezes, Peter S  
; APPLICANT: MacDougall, John  
; TITLE OF INVENTION: No. US2002015115A1el Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 15966-697 CIP  
; CURRENT APPLICATION NUMBER: US/09/808,602  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 09/800,198  
; PRIOR FILING DATE: 2001-03-05  
; PRIOR APPLICATION NUMBER: 60/186,596  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 106  
; LENGTH: 321  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-808-602-106

Query Match Best Local Similarity 100.0%; Score 21; DB 9; Length 321;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4

Db 208 SYDA 211

## RESULT 28

US-09-738-626-4308  
; Sequence 4308, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAOKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OKAZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; PRIOR FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: Patentln ver. 3.0  
; SEQ ID NO 4308  
; LENGTH: 332  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-4308

Query Match 100.0%; Score 21; DB 9; Length 332;

Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4

Db 314 SYDA 317

## RESULT 29

US-09-962-618-2  
; Sequence 2, Application US/09962618  
; Patent No. US2002010357A1  
; GENERAL INFORMATION:  
; APPLICANT: BATHE, Brigitte  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE msik GENE  
; FILE REFERENCE: 032301 WD 228  
; CURRENT APPLICATION NUMBER: US/09/962,618  
; PRIOR FILING DATE: 2001-01-19  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: Patentln version 3.1  
; SEQ ID NO 2  
; LENGTH: 332  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-09-962-618-2

Query Match Best Local Similarity 100.0%; Score 21; DB 10; Length 332;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4

Db 314 SYDA 317

## RESULT 30

US-09-815-242-10953  
; Sequence 10953, Application US/09815242  
; Patent No. US2002061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes In  
; FILE REFERENCE: ELITRA-011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10953  
; LENGTH: 339  
; TYPE: PRT  
; ORGANISM: Haemophilus influenzae

US-09-815-242-10953

Query Match 100.0%; Score 21; DB 10; Length 339;

Best Local Similarity 100.0%; Pred. No. 3.8e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4

1111

Db 252 SYDA 255

RESULT 31

US-09-815-242-10844

; Sequence 10844, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Karl L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 10844

; LENGTH: 343

; TYPE: PRT

; ORGANISM: Enterococcus faecalis

; US-09-815-242-10844

Query Match 100.0%; Score 21; DB 10; Length 343;

Best Local Similarity 100.0%; Pred. No. 3.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4

1111

Db 178 SYDA 181

RESULT 32

US-09-738-626-4252

; Sequence 4252, Application US/09738626

; Publication No. US20020197605A1

; GENERAL INFORMATION:

; APPLICANT: NAKAGAWA, SATOSHI

; APPLICANT: MIZOGUCHI, HIROSHI

; APPLICANT: ANDO, SEIKO

; APPLICANT: HAYASHI, MIKIRO

; APPLICANT: OCHIAI, KEIKO

; APPLICANT: YOKOI, HARUHIKO

; APPLICANT: TATEISHI, NAKO

; APPLICANT: SENOH, AKIHITO

; APPLICANT: IKEDA, MASATO

; APPLICANT: OZAKI, AKIO

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-125

; CURRENT APPLICATION NUMBER: US/09/738,626

; PRIOR FILING DATE: 2000-12-18

; PRIOR APPLICATION NUMBER: JP 99/377484

; PRIOR FILING DATE: 1999-12-16

; PRIOR APPLICATION NUMBER: JP 00/159162

; PRIOR FILING DATE: 2000-04-07

; PRIOR APPLICATION NUMBER: JP 00/280988

; PRIOR FILING DATE: 2000-08-03

; NUMBER OF SEQ ID NOS: 7059

; SOFTWARE: Patentln Ver. 3.0

; SEQ ID NO 4252

; LENGTH: 345

; TYPE: PRT

; ORGANISM: Corynebacterium glutamicum

; US-09-738-626-4252

Query Match 100.0%; Score 21; DB 9; Length 345;

Best Local Similarity 100.0%; Pred. No. 3.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4

1111

Db 45 SYDA 48

RESULT 33

US-09-808-602-105

; Sequence 105, Application US/09808602

; Patent No. US2002015115A1

; GENERAL INFORMATION:

; APPLICANT: Vernet, Corine A

; APPLICANT: Fernandes, Elma

; APPLICANT: Shimkets, Richard A

; APPLICANT: Herrman, John L

; APPLICANT: Majumder, Kumud

; APPLICANT: Mishra, Vishnu

; APPLICANT: Mezes, Peter S

; APPLICANT: MacDougall, John

; TITLE OF INVENTION: No. US2002015115A1 Proteins and Nucleic Acids Encoding Same

; FILE REFERENCE: 15966-697 CIP

; CURRENT APPLICATION NUMBER: US/09/808,602

; PRIOR FILING DATE: 2001-03-14

; PRIOR APPLICATION NUMBER: 09/800,198

; PRIOR FILING DATE: 2001-03-05

; PRIOR APPLICATION NUMBER: 60/186,596

; PRIOR FILING DATE: 2000-03-03

; NUMBER OF SEQ ID NOS: 114

; SOFTWARE: Patentln Ver. 2.1

; SEQ ID NO 105

; LENGTH: 352

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-808-602-105

Query Match 100.0%; Score 21; DB 9; Length 352;

Best Local Similarity 100.0%; Pred. No. 4e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4

1111

Db 242 SYDA 245

RESULT 34

US-09-888-615-95

; Sequence 95, Application US/09888615

; Patent No. US20020064856A1

; GENERAL INFORMATION:

; APPLICANT: PLOWMAN, GREGORY

APPLICANT: WHITE, DAVID  
APPLICANT: CAENEPEEL, SEAN  
APPLICANT: CHARVETZAK, GLEN  
APPLICANT: MANNING, GERARD  
APPLICANT: SUDARSANAM, SUCHA  
TITLE OF INVENTION: NOVEL PROTEASES  
FILE REFERENCE: 038602/1214  
CURRENT APPLICATION NUMBER: US/09/888,615  
CURRENT FILING DATE: 2001-06-26  
PRIOR APPLICATION NUMBER: 60/214,047  
PRIOR FILING DATE: 2000-06-26  
NUMBER OF SEQ ID NOS: 150  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 95  
LENGTH: 352  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-888-615-95

Query Match  
Best Local Similarity 100.0%; Score 21; DB 10; Length 352;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
Db 242 SYDA 245

RESULT 35  
US-09-799-777-45  
Sequence 45, Application US/09799777  
Patent No. US20020091244A1  
GENERAL INFORMATION:  
APPLICANT: Lal, Preeti  
Hillman, Jennifer L.  
Corley, Neil C.  
Guegler, Karl J.  
Baugh, Mariah  
Sather, Susan  
Shah, Purvi  
TITLE OF INVENTION: HUMAN SIGNAL PEPTIDE-CONTAINING PROTEINS  
NUMBER OF SEQUENCES: 154  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/799,777  
FILING DATE: 06-Mar-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/002,485  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: BILLINGS, LUCY J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0459 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 45:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 359 amino acids  
TYPE: amino acid  
STRANDEDNESS: single

TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: LUNCAST01  
CLONE: 1988911  
US-09-799-777-45  
SEQUENCE DESCRIPTION: SEQ ID NO: 45 :

Query Match  
Best Local Similarity 100.0%; Score 21; DB 10; Length 359;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
Db 194 SYDA 197

RESULT 36  
US-09-738-626-5241  
Sequence 5241, Application US/09738626  
Publication No. US20020197605A1  
GENERAL INFORMATION:  
APPLICANT: NAKAGAWA, SATOSHI  
APPLICANT: MIZOGUCHI, HIROSHI  
APPLICANT: ANDO, SEIKO  
APPLICANT: HAYASHI, MIKIRO  
APPLICANT: OCHIAI, KEIKO  
APPLICANT: YOKOI, HARUHIKO  
APPLICANT: TATEISHI, NAOKO  
APPLICANT: SENOH, AKIHIRO  
APPLICANT: IKEDA, MASATO  
APPLICANT: OZAKI, AKIO  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-125  
CURRENT APPLICATION NUMBER: US/09/738, 626  
CURRENT FILING DATE: 2000-12-18  
PRIOR APPLICATION NUMBER: JP 99/377484  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: JP 00/159162  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: JP 00/280988  
PRIOR FILING DATE: 2000-08-03  
NUMBER OF SEQ ID NOS: 7059  
SOFTWARE: PatentIn ver. 3.0  
SEQ ID NO 5241  
LENGTH: 360  
TYPE: PRT  
ORGANISM: Corynebacterium glutamicum  
US-09-738-626-5241

Query Match  
Best Local Similarity 100.0%; Score 21; DB 9; Length 360;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
Db 57 SYDA 60

RESULT 37  
US-09-815-242-11378  
Sequence 11378, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes In Prokaryotes

```
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11378
LENGTH: 360
TYPE: PRT
ORGANISM: Helicobacter pylori
US-09-815-242-11378
```

```
Query Match
Best Local Similarity 100.0%; Score 21; DB 10; Length 360;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYDA 4
||||
DB 228 SYDA 231
```

```
RESULT 38
US-09-738-626-5738
Sequence 5738, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 5738
LENGTH: 378
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-5738
```

```
Query Match
Best Local Similarity 100.0%; Score 21; DB 9; Length 378;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYDA 4
||||
```

```
DB 249 SYDA 252
```

```
RESULT 39
US-09-971-536-60
Sequence 60, Application US/09971536
Patent No. US20020159976A1
GENERAL INFORMATION:
APPLICANT: Glenn, Matthew
APPLICANT: Havukkala, Ilkka
APPLICANT: Bloksberg, Leonard
APPLICANT: Lubbers, Mark
APPLICANT: Dekker, James
APPLICANT: Christensson, Anna
APPLICANT: Holland, Ross
APPLICANT: O'Toole, Paul
APPLICANT: Reid, Julian
APPLICANT: Coolbear, Timothy
TITLE OF INVENTION: Lactobacillus rhamnosus Polynucleotides, Polypeptides and Meth
TITLE OF INVENTION: Using Them
FILE REFERENCE: 104302
CURRENT APPLICATION NUMBER: US/09/971,536
CURRENT FILING DATE: 2001-10-02
PRIOR APPLICATION NUMBER: U.S. No. US20020159976A1 09/634,238
PRIOR FILING DATE: 2000-08-08
PRIOR APPLICATION NUMBER: U.S. No. US20020159976A1 09/724,623
PRIOR FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: PCT/NZ01/00160
NUMBER OF SEQ ID NOS: 83
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 60
LENGTH: 416
TYPE: PRT
ORGANISM: Lactobacillus rhamnosus
US-09-971-536-60
```

```
Query Match
Best Local Similarity 100.0%; Score 21; DB 9; Length 416;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYDA 4
||||
DB 281 SYDA 284
```

```
RESULT 40
US-09-815-242-5696
Sequence 5696, Application US/09815242
Patent No. US2002061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
```



```
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5696
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5696

Query Match
Best Local Similarity 100.0%; Score 21; DB 10; Length 439;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
    ||||
Db 198 SYDA 201

RESULT 41
US-09-820-893-56
; Sequence 56, Application US/09820893
; Patent No. US20020076705A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 31 Human Secreted Proteins
; FILE REFERENCE: P2033P1
; CURRENT APPLICATION NUMBER: US/09/820,893
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/531,119
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/102,895
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 56
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-820-893-56

Query Match
Best Local Similarity 100.0%; Score 21; DB 10; Length 491;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
    ||||
Db 81 SYDA 84

RESULT 42
US-09-815-242-5045
; Sequence 5045, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
```

```
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5045
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-815-242-5045

Query Match
Best Local Similarity 100.0%; Score 21; DB 10; Length 506;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
    ||||
Db 165 SYDA 168

RESULT 43
US-09-815-242-11757
; Sequence 11757, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11757
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-815-242-11757

Query Match
Best Local Similarity 100.0%; Score 21; DB 10; Length 506;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 SYDA 4  
 ||||  
 Db 165 SYDA 168

RESULT 44  
 US-09-925-302-477

; Sequence 477, Application US/09925302  
 ; Patent No. US20020044941A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
 ; FILE REFERENCE: PA104  
 ; CURRENT APPLICATION NUMBER: US/09/925,302  
 ; CURRENT FILING DATE: 2001-08-10  
 ; PRIOR APPLICATION NUMBER: PCT/US00/05918  
 ; PRIOR FILING DATE: 2000-03-08  
 ; PRIOR APPLICATION NUMBER: 60/124,270  
 ; PRIOR FILING DATE: 1999-03-12  
 ; NUMBER OF SEQ ID NOS: 896  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 477  
 ; LENGTH: 549  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: SITE  
 ; LOCATION: (217)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; NAME/KEY: SITE  
 ; LOCATION: (224)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; US-09-925-302-477

Query Match

Best Local Similarity 100.0%; Score 21; DB 10; Length 549;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
 ||||  
 Db 225 SYDA 228

RESULT 45

US-09-824-735-4

; Sequence 4, Application US/09824735  
 ; Patent No. US20020095032A1

; GENERAL INFORMATION:

; APPLICANT: ZHU, JIAN-KANG

; APPLICANT: LIU, JIPING

; APPLICANT: ISHITANI, MANABU

; APPLICANT: HALFTER, URSULA

; APPLICANT: KIM, CHEOL-SOO

; TITLE OF INVENTION: PROTEINS AND DNA RELATED TO SALT TOLERANCE IN PLANTS

; FILE REFERENCE: 205645US20

; CURRENT APPLICATION NUMBER: US/09/824,735

; CURRENT FILING DATE: 2001-08-17

; PRIOR APPLICATION NUMBER: US 60/824,735

; PRIOR FILING DATE: 2000-04-04

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4

; LENGTH: 552

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-824-735-4

Query Match 100.0%; Score 21; DB 10; Length 552;

Best Local Similarity 100.0%; Pred. No. 6.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
 ||||

Db 282 SYDA 285

Search completed: February 6, 2003, 11:39:44  
 Job time : 6.83333 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model1

Run on: February 6, 2003, 11:16:59 : Search time 133.167 Seconds  
(without alignments)  
19.366 Million cell updates/sec

Title: PAT943-3  
Perfect score: 21  
Sequence: 1 syda 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 64473110 residues  
Total number of hits satisfying chosen parameters: 3388

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%  
Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/1/paa/PCrUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US080\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/1/paa/US097\_COMB.pep.\*  
22: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*  
23: /cgn2\_6/ptodata/1/paa/US099\_COMB.pep.\*  
24: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*  
25: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*  
26: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*  
27: /cgn2\_6/ptodata/1/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	100.0	5	23	US-09-908-943A-178
2	21	100.0	5	27	US-60-275-251-178
3	21	100.0	7	7	US-08-310-912-126
4	21	100.0	7	22	US-09-867-852-126
5	21	100.0	9	1	PCT-US02-11643-95
6	21	100.0	9	1	PCT-US02-11643-307

7	21	100.0	9	1	PCT-US02-11643-552	Sequence 552, App
8	21	100.0	9	1	PCT-US02-11643-556	Sequence 556, App
9	21	100.0	9	1	PCT-US02-11643-771	Sequence 771, App
10	21	100.0	9	1	PCT-US02-11643-817	Sequence 817, App
11	21	100.0	9	1	PCT-US02-11643-1043	Sequence 1043, App
12	21	100.0	9	1	PCT-US02-11643-1569	Sequence 1569, App
13	21	100.0	9	1	PCT-US02-11643-1760	Sequence 1760, App
14	21	100.0	9	1	PCT-US02-11643-1838	Sequence 1838, App
15	21	100.0	9	1	PCT-US02-11643-1912	Sequence 1912, App
16	21	100.0	9	1	PCT-US02-11643-1968	Sequence 1968, App
17	21	100.0	9	1	PCT-US02-11643-2132	Sequence 2132, App
18	21	100.0	9	1	PCT-US02-11643-2379	Sequence 2379, App
19	21	100.0	9	1	PCT-US02-11643-2398	Sequence 2398, App
20	21	100.0	9	1	PCT-US02-11643-2579	Sequence 2579, App
21	21	100.0	9	1	PCT-US02-11643-2733	Sequence 2733, App
22	21	100.0	9	1	PCT-US02-11643-2862	Sequence 2862, App
23	21	100.0	9	1	PCT-US02-11643-2863	Sequence 2863, App
24	21	100.0	9	1	PCT-US02-11643-3160	Sequence 3160, App
25	21	100.0	9	1	PCT-US02-11643-3161	Sequence 3161, App
26	21	100.0	9	1	PCT-US02-11643-3188	Sequence 3188, App
27	21	100.0	9	1	PCT-US02-11643-3305	Sequence 3305, App
28	21	100.0	9	1	PCT-US02-11643-3306	Sequence 3306, App
29	21	100.0	9	1	PCT-US02-11643-3496	Sequence 3496, App
30	21	100.0	9	1	PCT-US02-11643-3758	Sequence 3758, App
31	21	100.0	9	1	PCT-US02-11643-3910	Sequence 3910, App
32	21	100.0	9	1	PCT-US02-11643-3971	Sequence 3971, App
33	21	100.0	9	1	PCT-US02-11643-4066	Sequence 4066, App
34	21	100.0	9	1	PCT-US02-11643-4149	Sequence 4149, App
35	21	100.0	9	1	PCT-US02-11643-4204	Sequence 4204, App
36	21	100.0	9	1	PCT-US02-11643-4385	Sequence 4385, App
37	21	100.0	9	1	PCT-US02-11643-4448	Sequence 4448, App
38	21	100.0	9	1	PCT-US02-11643-4532	Sequence 4532, App
39	21	100.0	9	1	PCT-US02-11643-4647	Sequence 4647, App
40	21	100.0	9	1	PCT-US02-11643-4799	Sequence 4799, App
41	21	100.0	9	1	PCT-US02-11643-4800	Sequence 4800, App
42	21	100.0	9	1	PCT-US02-11643-4949	Sequence 4949, App
43	21	100.0	9	1	PCT-US02-11643-5109	Sequence 5109, App
44	21	100.0	9	1	PCT-US02-11643-5244	Sequence 5244, App
45	21	100.0	9	1	PCT-US02-11643-5290	Sequence 5290, App

#### ALIGNMENTS

RESULT 1  
US-09-908-943A-178  
Sequence 178, Application US/09908943A  
GENERAL INFORMATION:  
APPLICANT: Yan, Riqiang  
APPLICANT: Tomasselli, Alfredo G.  
APPLICANT: Guirney, Mark E.  
APPLICANT: Emmons, Thomas L.  
APPLICANT: Bienkowski, Mike J.  
APPLICANT: Helirison, Robert L.  
TITLE OF INVENTION: SUBSTRATES AND ASSAYS FOR BETA-SECRETASE ACTIVITY  
FILE REFERENCE: 29915/00281A.US1  
CURRENT APPLICATION NUMBER: US/09/908, 943A  
CURRENT FILING DATE: 2001-07-19  
PRIOR APPLICATION NUMBER: 60/219,795  
PRIOR FILING DATE: 2000-07-19  
NUMBER OF SEQ ID NOS: 197  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 178  
LENGTH: 5  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-908-943A-178  
Query Match 100.0% Score 21: DB 23: Length 5:  
Best local Similarity 100.0% Pred. No. 4.2e+06;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
|||||  
Db 2 SYDA 5

RESULT 2  
US-60-275-251-178  
; Sequence 178, Application US/60275251  
; GENERAL INFORMATION:  
; APPLICANT: Yan, Ridiang  
; APPLICANT: Tomasselli, Alfredo G.  
; APPLICANT: Gurney, Mark E.  
; APPLICANT: Emmons, Thomas L.  
; APPLICANT: Bienkowski, Mike J.  
; APPLICANT: Helirison, Robert L.  
; TITLE OF INVENTION: SUBSTRATES AND ASSAYS FOR BETA-SECRETASE ACTIVITY  
; FILE REFERENCE: 29915/00281  
; CURRENT APPLICATION NUMBER: US/60/275,251  
; NUMBER OF SEQ ID NOS: 197  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 178  
; LENGTH: 5  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: peptide sequence  
US-60-275-251-178

Query Match 100.0%; Score 21; DB 27; Length 5;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
|||||  
Db 2 SYDA 5

RESULT 3  
US-08-310-912-126  
; Sequence 126, Application US/08310912  
; GENERAL INFORMATION:  
; APPLICANT: Ausubel, Frederick M.  
; APPLICANT: Staskawicz, Brian J.  
; APPLICANT: Brent, Andrew F.  
; APPLICANT: Dahlbeck, Douglas  
; APPLICANT: Katagiri, Fumiaki  
; APPLICANT: Kunkel, Barbara N.  
; APPLICANT: Mindinos, Michael N.  
; APPLICANT: Yu, Guo-Liang  
; TITLE OF INVENTION: RPS2 GENE AND USES THEREOF  
; NUMBER OF SEQUENCES: 201  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street Suite 3100  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2904  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30B  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/310,912  
; FILING DATE:  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/227,360

; FILING DATE: 13-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clark, Paul T.  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 00786/230001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 100254  
; INFORMATION FOR SEQ ID NO: 126:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-310-912-126

Query Match 100.0%; Score 21; DB 7; Length 7;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
|||||  
Db 3 SYDA 6

RESULT 4  
US-09-867-852-126  
; Sequence 126, Application US/09867852  
; GENERAL INFORMATION:  
; APPLICANT: Ausubel, Frederick M.  
; APPLICANT: Staskawicz, Brian J.  
; APPLICANT: Brent, Andrew F.  
; APPLICANT: Dahlbeck, Douglas  
; APPLICANT: Katagiri, Fumiaki  
; APPLICANT: Kunkel, Barbara N.  
; APPLICANT: Mindinos, Michael N.  
; APPLICANT: Yu, Guo-Liang  
; TITLE OF INVENTION: RPS2 GENE FAMILY, PRIMERS, PROBES, AND  
; FILE REFERENCE: 00786/254002  
; CURRENT APPLICATION NUMBER: US/09/867,852  
; CURRENT FILING DATE: 2001-05-29  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/301,085  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-28  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/310,912  
; PRIOR FILING DATE: EARLIER FILING DATE: 1994-09-22  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/227,360  
; PRIOR FILING DATE: EARLIER FILING DATE: 1994-04-13  
; NUMBER OF SEQ ID NOS: 208  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 126  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-09-867-852-126

Query Match 100.0%; Score 21; DB 22; Length 7;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
|||||  
Db 3 SYDA 6

RESULT 5  
PCT-US02-11643-95  
; Sequence 95, Application PC/TUS0211643  
; GENERAL INFORMATION:  
; APPLICANT: Chailita-Eid, Pia M.  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Farris, Mary

```
; APPLICANT: Rubert, Rene S.
; APPLICANT: Morrison, Karen J.
; APPLICANT: Morrison, Robert K.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20067.40
; CURRENT APPLICATION NUMBER: PCT/US02/11643
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60,286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 15017
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 95
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapiens
PCT-US02-11643-95
```

```
Query Match          100.0%; Score 21; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYDA 4
    ||||
Db 4 SYDA 7
```

```
RESULT 5
PCT-US02-11643-307
; Sequence 307, Application PC/TUS0211643
; GENERAL INFORMATION:
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Farris, Mary
; APPLICANT: Rubert, Rene S.
; APPLICANT: Morrison, Karen J.
; APPLICANT: Morrison, Robert K.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20067.40
; CURRENT APPLICATION NUMBER: PCT/US02/11643
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60,286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 15017
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 307
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapiens
PCT-US02-11643-307
```

```
Query Match          100.0%; Score 21; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYDA 4
    ||||
Db 6 SYDA 9
```

```
RESULT 7
PCT-US02-11643-552
; Sequence 552, Application PC/TUS0211643
; GENERAL INFORMATION:
```

```
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Farris, Mary
; APPLICANT: Rubert, Rene S.
; APPLICANT: Morrison, Karen J.
; APPLICANT: Morrison, Robert K.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20067.40
; CURRENT APPLICATION NUMBER: PCT/US02/11643
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60,286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 15017
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 552
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapiens
PCT-US02-11643-552
```

```
Query Match          100.0%; Score 21; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYDA 4
    ||||
Db 6 SYDA 9
```

```
RESULT 8
PCT-US02-11643-556
; Sequence 556, Application PC/TUS0211643
; GENERAL INFORMATION:
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Farris, Mary
; APPLICANT: Rubert, Rene S.
; APPLICANT: Morrison, Karen J.
; APPLICANT: Morrison, Robert K.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20067.40
; CURRENT APPLICATION NUMBER: PCT/US02/11643
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60,286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 15017
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 556
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapiens
PCT-US02-11643-556
```

```
Query Match          100.0%; Score 21; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYDA 4
    ||||
Db 4 SYDA 7
```

```
RESULT 9
```

```
PCT-US02-11643-771
; Sequence 771, Application PC/TUS0211643
; GENERAL INFORMATION:
; APPLICANT: Challita-Eld, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Rubert, Rene S.
; APPLICANT: Morrison, Karen J.
; APPLICANT: Morrison, Robert K.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20067.40
; CURRENT APPLICATION NUMBER: PCT/US02/11643
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60,286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 15017
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 771
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapiens
PCT-US02-11643-771
```

```
Query Match          100.0%; Score 21; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYDA 4
    ||||
DB 4 SYDA 7
```

```
RESULT 10
PCT-US02-11643-817
; Sequence 817, Application PC/TUS0211643
; GENERAL INFORMATION:
; APPLICANT: Challita-Eld, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Rubert, Rene S.
; APPLICANT: Morrison, Karen J.
; APPLICANT: Morrison, Robert K.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20067.40
; CURRENT APPLICATION NUMBER: PCT/US02/11643
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60,286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 15017
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 817
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapiens
PCT-US02-11643-817
```

```
Query Match          100.0%; Score 21; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYDA 4
    ||||
DB 5 SYDA 8
```

```
RESULT 11
PCT-US02-11643-1043
; Sequence 1043, Application PC/TUS0211643
; GENERAL INFORMATION:
; APPLICANT: Challita-Eld, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Rubert, Rene S.
; APPLICANT: Morrison, Karen J.
; APPLICANT: Morrison, Robert K.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20067.40
; CURRENT APPLICATION NUMBER: PCT/US02/11643
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60,286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 15017
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1043
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapiens
PCT-US02-11643-1043
```

```
Query Match          100.0%; Score 21; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYDA 4
    ||||
DB 1 SYDA 4
```

```
RESULT 12
PCT-US02-11643-1569
; Sequence 1569, Application PC/TUS0211643
; GENERAL INFORMATION:
; APPLICANT: Challita-Eld, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Rubert, Rene S.
; APPLICANT: Morrison, Karen J.
; APPLICANT: Morrison, Robert K.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20067.40
; CURRENT APPLICATION NUMBER: PCT/US02/11643
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60,286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 15017
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1569
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapiens
PCT-US02-11643-1569
```

```
Query Match          100.0%; Score 21; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

OY 1 SYDA 4  
1111  
Db 3 SYDA 6

RESULT 13  
PCT-US02-11643-1760

; Sequence 1760, Application PC/TUS0211643  
; GENERAL INFORMATION:  
; APPLICANT: Chaillita-Elid, Pia M.  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Faris, Mary  
; APPLICANT: Rubert, Rene S.  
; APPLICANT: Morrison, Karen J.  
; APPLICANT: Morrison, Robert K.  
; APPLICANT: Ge, Wangmao  
; APPLICANT: Jakobovits, Aya  
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN  
; FILE REFERENCE: 51158-20067.40  
; CURRENT APPLICATION NUMBER: PCT/US02/11643  
; CURRENT FILING DATE: 2002-04-09  
; PRIOR APPLICATION NUMBER: US 60/282,739  
; PRIOR FILING DATE: 2001-04-10  
; PRIOR APPLICATION NUMBER: US 60,286,630  
; PRIOR FILING DATE: 2001-04-25  
; NUMBER OF SEQ ID NOS: 15017  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1760  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
PCT-US02-11643-1760

Query Match 100.0%; Score 21; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 SYDA 4  
1111  
Db 1 SYDA 4

RESULT 14

PCT-US02-11643-1838

; Sequence 1838, Application PC/TUS0211643  
; GENERAL INFORMATION:  
; APPLICANT: Chaillita-Elid, Pia M.  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Faris, Mary  
; APPLICANT: Rubert, Rene S.  
; APPLICANT: Morrison, Karen J.  
; APPLICANT: Morrison, Robert K.  
; APPLICANT: Ge, Wangmao  
; APPLICANT: Jakobovits, Aya  
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN  
; FILE REFERENCE: 51158-20067.40  
; CURRENT APPLICATION NUMBER: PCT/US02/11643  
; CURRENT FILING DATE: 2002-04-09  
; PRIOR APPLICATION NUMBER: US 60/282,739  
; PRIOR FILING DATE: 2001-04-10  
; PRIOR APPLICATION NUMBER: US 60,286,630  
; PRIOR FILING DATE: 2001-04-25  
; NUMBER OF SEQ ID NOS: 15017  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1838  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
PCT-US02-11643-1838

Query Match 100.0%; Score 21; DB 1; Length 9;

Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
1111  
Db 4 SYDA 7

RESULT 15  
PCT-US02-11643-1912

; Sequence 1912, Application PC/TUS0211643  
; GENERAL INFORMATION:  
; APPLICANT: Chaillita-Elid, Pia M.  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Faris, Mary  
; APPLICANT: Rubert, Rene S.  
; APPLICANT: Morrison, Karen J.  
; APPLICANT: Morrison, Robert K.  
; APPLICANT: Ge, Wangmao  
; APPLICANT: Jakobovits, Aya  
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN  
; FILE REFERENCE: 51158-20067.40  
; CURRENT APPLICATION NUMBER: PCT/US02/11643  
; CURRENT FILING DATE: 2002-04-09  
; PRIOR APPLICATION NUMBER: US 60/282,739  
; PRIOR FILING DATE: 2001-04-10  
; PRIOR APPLICATION NUMBER: US 60,286,630  
; PRIOR FILING DATE: 2001-04-25  
; NUMBER OF SEQ ID NOS: 15017  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1912  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
PCT-US02-11643-1912

Query Match 100.0%; Score 21; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
1111  
Db 2 SYDA 5

RESULT 16  
PCT-US02-11643-1988

; Sequence 1988, Application PC/TUS0211643  
; GENERAL INFORMATION:  
; APPLICANT: Chaillita-Elid, Pia M.  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Faris, Mary  
; APPLICANT: Rubert, Rene S.  
; APPLICANT: Morrison, Karen J.  
; APPLICANT: Morrison, Robert K.  
; APPLICANT: Ge, Wangmao  
; APPLICANT: Jakobovits, Aya  
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN  
; FILE REFERENCE: 51158-20067.40  
; CURRENT APPLICATION NUMBER: PCT/US02/11643  
; CURRENT FILING DATE: 2002-04-09  
; PRIOR APPLICATION NUMBER: US 60/282,739  
; PRIOR FILING DATE: 2001-04-10  
; PRIOR APPLICATION NUMBER: US 60,286,630  
; PRIOR FILING DATE: 2001-04-25  
; NUMBER OF SEQ ID NOS: 15017  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1988  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo Sapiens

PCT-US02-11643-1988

Query Match 100.0%; Score 21; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
1111  
Db 3 SYDA 6

RESULT 17

PCT-US02-11643-2132  
; Sequence 2132, Application PC/TUS0211643  
; GENERAL INFORMATION:  
; APPLICANT: Challita-Eld, Pia M.  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Faris, Mary  
; APPLICANT: Rubert, Rene S.  
; APPLICANT: Morrison, Karen J.  
; APPLICANT: Morrison, Robert K.  
; APPLICANT: Ge, Wangmao  
; APPLICANT: Jakobovits, Aya  
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN  
; FILE REFERENCE: 51158-20067.40  
; CURRENT APPLICATION NUMBER: PCT/US02/11643  
; CURRENT FILING DATE: 2002-04-09  
; PRIOR APPLICATION NUMBER: US 60/286,739  
; PRIOR FILING DATE: 2001-04-10  
; PRIOR APPLICATION NUMBER: US 60,286,630  
; PRIOR FILING DATE: 2001-04-25  
; NUMBER OF SEQ ID NOS: 15017  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2132  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
PCT-US02-11643-2132

Query Match 100.0%; Score 21; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
1111  
Db 6 SYDA 9

RESULT 18

PCT-US02-11643-2379  
; Sequence 2379, Application PC/TUS0211643  
; GENERAL INFORMATION:  
; APPLICANT: Challita-Eld, Pia M.  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Faris, Mary  
; APPLICANT: Rubert, Rene S.  
; APPLICANT: Morrison, Karen J.  
; APPLICANT: Morrison, Robert K.  
; APPLICANT: Ge, Wangmao  
; APPLICANT: Jakobovits, Aya  
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN  
; FILE REFERENCE: 51158-20067.40  
; CURRENT APPLICATION NUMBER: PCT/US02/11643  
; CURRENT FILING DATE: 2002-04-09  
; PRIOR APPLICATION NUMBER: US 60/282,739  
; PRIOR FILING DATE: 2001-04-10  
; PRIOR APPLICATION NUMBER: US 60,286,630  
; PRIOR FILING DATE: 2001-04-25  
; NUMBER OF SEQ ID NOS: 15017  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2379

; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
PCT-US02-11643-2379

Query Match 100.0%; Score 21; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
1111  
Db 6 SYDA 9

RESULT 19

PCT-US02-11643-2398  
; Sequence 2398, Application PC/TUS0211643  
; GENERAL INFORMATION:  
; APPLICANT: Challita-Eld, Pia M.  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Faris, Mary  
; APPLICANT: Rubert, Rene S.  
; APPLICANT: Morrison, Karen J.  
; APPLICANT: Morrison, Robert K.  
; APPLICANT: Ge, Wangmao  
; APPLICANT: Jakobovits, Aya  
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN  
; FILE REFERENCE: 51158-20067.40  
; CURRENT APPLICATION NUMBER: PCT/US02/11643  
; CURRENT FILING DATE: 2002-04-09  
; PRIOR APPLICATION NUMBER: US 60/282,739  
; PRIOR FILING DATE: 2001-04-10  
; PRIOR APPLICATION NUMBER: US 60,286,630  
; PRIOR FILING DATE: 2001-04-25  
; NUMBER OF SEQ ID NOS: 15017  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2398  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
PCT-US02-11643-2398

Query Match 100.0%; Score 21; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
1111  
Db 1 SYDA 4

RESULT 20

PCT-US02-11643-2579  
; Sequence 2579, Application PC/TUS0211643  
; GENERAL INFORMATION:  
; APPLICANT: Challita-Eld, Pia M.  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Faris, Mary  
; APPLICANT: Rubert, Rene S.  
; APPLICANT: Morrison, Karen J.  
; APPLICANT: Morrison, Robert K.  
; APPLICANT: Ge, Wangmao  
; APPLICANT: Jakobovits, Aya  
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN  
; FILE REFERENCE: 51158-20067.40  
; CURRENT APPLICATION NUMBER: PCT/US02/11643  
; CURRENT FILING DATE: 2002-04-09  
; PRIOR APPLICATION NUMBER: US 60/282,739  
; PRIOR FILING DATE: 2001-04-10  
; PRIOR APPLICATION NUMBER: US 60,286,630  
; PRIOR FILING DATE: 2001-04-25



```
; NUMBER OF SEQ ID NOS: 15017
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2579
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapiens
PCT-US02-11643-2579
```

```
Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYDA 4
    ||||
Db 4 SYDA 7
```

```
RESULT 21
PCT-US02-11643-2733
```

```
; Sequence 2733, Application PC/TUS0211643
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Challita-Eid, Pia M.
```

```
; APPLICANT: Raitano, Arthur B.
```

```
; APPLICANT: Faris, Mary
```

```
; APPLICANT: Rubert, Rene S.
```

```
; APPLICANT: Morrison, Karen J.
```

```
; APPLICANT: Morrison, Robert K.
```

```
; APPLICANT: Ge, Wangmao
```

```
; APPLICANT: Jakobovits, Aya
```

```
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
```

```
; FILE REFERENCE: ENTITLED 184PIE2 USEFUL IN TREATMENT AND DETECTION OF CANCER
```

```
; CURRENT FILING DATE: 2002-04-09
```

```
; PRIOR FILING DATE: 2001-04-10
```

```
; PRIOR APPLICATION NUMBER: US 60,286,630
```

```
; PRIOR FILING DATE: 2001-04-25
```

```
; NUMBER OF SEQ ID NOS: 15017
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 2733
```

```
; LENGTH: 9
```

```
; TYPE: PRT
```

```
; ORGANISM: Homo Sapiens
```

```
PCT-US02-11643-2733
```

```
Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYDA 4
    ||||
Db 2 SYDA 5
```

```
RESULT 22
PCT-US02-11643-2862
```

```
; Sequence 2862, Application PC/TUS0211643
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Challita-Eid, Pia M.
```

```
; APPLICANT: Raitano, Arthur B.
```

```
; APPLICANT: Paris, Mary
```

```
; APPLICANT: Rubert, Rene S.
```

```
; APPLICANT: Morrison, Karen J.
```

```
; APPLICANT: Morrison, Robert K.
```

```
; APPLICANT: Ge, Wangmao
```

```
; APPLICANT: Jakobovits, Aya
```

```
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
```

```
; FILE REFERENCE: ENTITLED 184PIE2 USEFUL IN TREATMENT AND DETECTION OF CANCER
```

```
; CURRENT FILING DATE: 2002-04-09
```

```
; PRIOR APPLICATION NUMBER: US 60,282,739
```

```
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60,286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 15017
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2862
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapiens
PCT-US02-11643-2862
```

```
Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYDA 4
    ||||
Db 5 SYDA 8
```

```
RESULT 23
PCT-US02-11643-2863
```

```
; Sequence 2863, Application PC/TUS0211643
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Challita-Eid, Pia M.
```

```
; APPLICANT: Raitano, Arthur B.
```

```
; APPLICANT: Faris, Mary
```

```
; APPLICANT: Rubert, Rene S.
```

```
; APPLICANT: Morrison, Karen J.
```

```
; APPLICANT: Morrison, Robert K.
```

```
; APPLICANT: Ge, Wangmao
```

```
; APPLICANT: Jakobovits, Aya
```

```
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
```

```
; FILE REFERENCE: ENTITLED 184PIE2 USEFUL IN TREATMENT AND DETECTION OF CANCER
```

```
; CURRENT FILING DATE: 2002-04-09
```

```
; PRIOR FILING DATE: 2001-04-10
```

```
; PRIOR APPLICATION NUMBER: US 60,282,739
```

```
; PRIOR FILING DATE: 2001-04-25
```

```
; NUMBER OF SEQ ID NOS: 15017
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 2863
```

```
; LENGTH: 9
```

```
; TYPE: PRT
```

```
; ORGANISM: Homo Sapiens
```

```
PCT-US02-11643-2863
```

```
Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYDA 4
    ||||
Db 3 SYDA 6
```

```
RESULT 24
PCT-US02-11643-3160
```

```
; Sequence 3160, Application PC/TUS0211643
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Challita-Eid, Pia M.
```

```
; APPLICANT: Raitano, Arthur B.
```

```
; APPLICANT: Faris, Mary
```

```
; APPLICANT: Rubert, Rene S.
```

```
; APPLICANT: Morrison, Karen J.
```

```
; APPLICANT: Morrison, Robert K.
```

```
; APPLICANT: Ge, Wangmao
```

```
; APPLICANT: Jakobovits, Aya
```

```
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
```

```
; FILE REFERENCE: ENTITLED 184PIE2 USEFUL IN TREATMENT AND DETECTION OF CANCER
```

```

; CURRENT APPLICATION NUMBER: PCT/US02/11643
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60,286,630
; PRIOR FILING DATE: 2001-04-25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3160
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapiens
PCT-US02-11643-3160

```

```

Query Match          100.0%; Score 21; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SYDA 4
   ||||
Db 5 SYDA 8

```

```

RESULT 25
PCT-US02-11643-3161
; Sequence 3161, Application PC/TUS0211643
; GENERAL INFORMATION:
; APPLICANT: Chaillita-Eld, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Rubert, Rene S.
; APPLICANT: Morrison, Karen J.
; APPLICANT: Morrison, Robert K.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: ENTITLED 184PIE2 USEFUL IN TREATMENT AND DETECTION OF CANCER
; CURRENT APPLICATION NUMBER: PCT/US02/11643
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60,286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 15017
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3161
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapiens
PCT-US02-11643-3161

```

```

Query Match          100.0%; Score 21; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SYDA 4
   ||||
Db 1 SYDA 4

```

```

RESULT 26
PCT-US02-11643-3188
; Sequence 3188, Application PC/TUS0211643
; GENERAL INFORMATION:
; APPLICANT: Chaillita-Eld, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Rubert, Rene S.
; APPLICANT: Morrison, Karen J.
; APPLICANT: Morrison, Robert K.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya

```

```

; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: ENTITLED 184PIE2 USEFUL IN TREATMENT AND DETECTION OF CANCER
; CURRENT APPLICATION NUMBER: PCT/US02/11643
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60,286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 15017
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3188
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapiens
PCT-US02-11643-3188

```

```

Query Match          100.0%; Score 21; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SYDA 4
   ||||
Db 6 SYDA 9

```

```

RESULT 27
PCT-US02-11643-3305
; Sequence 3305, Application PC/TUS0211643
; GENERAL INFORMATION:
; APPLICANT: Chaillita-Eld, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Rubert, Rene S.
; APPLICANT: Morrison, Karen J.
; APPLICANT: Morrison, Robert K.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: ENTITLED 184PIE2 USEFUL IN TREATMENT AND DETECTION OF CANCER
; CURRENT APPLICATION NUMBER: PCT/US02/11643
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60,286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 15017
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3305
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapiens
PCT-US02-11643-3305

```

```

Query Match          100.0%; Score 21; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SYDA 4
   ||||
Db 4 SYDA 7

```

```

RESULT 28
PCT-US02-11643-3306
; Sequence 3306, Application PC/TUS0211643
; GENERAL INFORMATION:
; APPLICANT: Chaillita-Eld, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Rubert, Rene S.
; APPLICANT: Morrison, Karen J.

```

```

; APPLICANT: Morrison, Robert K.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20067.40
; CURRENT APPLICATION NUMBER: PCT/US02/11643
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60,286,630
; NUMBER OF SEQ ID NOS: 15017
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3306
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapiens
PCT-US02-11643-3306

```

```

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 1 SYDA 4
    |||
Db 2 SYDA 5

```

```

RESULT 29
PCT-US02-11643-3496
; Sequence 3496, Application PC/TUS0211643
; GENERAL INFORMATION:
; APPLICANT: Challita-Eld, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Rubert, Rene S.
; APPLICANT: Morrison, Karen J.
; APPLICANT: Morrison, Robert K.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20067.40
; CURRENT APPLICATION NUMBER: PCT/US02/11643
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60,286,630
; NUMBER OF SEQ ID NOS: 15017
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3496
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapiens
PCT-US02-11643-3496

```

```

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 1 SYDA 4
    |||
Db 3 SYDA 6

```

```

RESULT 30
PCT-US02-11643-3758
; Sequence 3758, Application PC/TUS0211643
; GENERAL INFORMATION:
; APPLICANT: Challita-Eld, Pia M.
; APPLICANT: Raitano, Arthur B.

```

```

; APPLICANT: Faris, Mary
; APPLICANT: Rubert, Rene S.
; APPLICANT: Morrison, Karen J.
; APPLICANT: Morrison, Robert K.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20067.40
; CURRENT APPLICATION NUMBER: PCT/US02/11643
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60,286,630
; NUMBER OF SEQ ID NOS: 15017
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3758
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapiens
PCT-US02-11643-3758

```

```

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 1 SYDA 4
    |||
Db 4 SYDA 7

```

```

RESULT 31
PCT-US02-11643-3910
; Sequence 3910, Application PC/TUS0211643
; GENERAL INFORMATION:
; APPLICANT: Challita-Eld, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Rubert, Rene S.
; APPLICANT: Morrison, Karen J.
; APPLICANT: Morrison, Robert K.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20067.40
; CURRENT APPLICATION NUMBER: PCT/US02/11643
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60,286,630
; NUMBER OF SEQ ID NOS: 15017
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3910
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapiens
PCT-US02-11643-3910

```

```

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 1 SYDA 4
    |||
Db 6 SYDA 9

```

```

RESULT 32
PCT-US02-11643-3971
; Sequence 3971, Application PC/TUS0211643

```

```

; GENERAL INFORMATION:
; APPLICANT: Challita-Eld, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Rubert, Rene S.
; APPLICANT: Morrison, Karen J.
; APPLICANT: Morrison, Robert K.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20067.40
; CURRENT APPLICATION NUMBER: PCT/US02/11643
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60,286,630
; NUMBER OF SEQ ID NOS: 15017
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3971
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapiens
PCT-US02-11643-3971

```

```

Query Match          100.0%; Score 21; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
OY 1 SYDA 4
    ||||
DB 5 SYDA 8
```

```

RESULT 33
PCT-US02-11643-4066
; Sequence 4066, Application PC/TUS0211643
; GENERAL INFORMATION:
; APPLICANT: Challita-Eld, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Rubert, Rene S.
; APPLICANT: Morrison, Karen J.
; APPLICANT: Morrison, Robert K.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20067.40
; CURRENT APPLICATION NUMBER: PCT/US02/11643
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60,286,630
; NUMBER OF SEQ ID NOS: 15017
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4066
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapiens
PCT-US02-11643-4066

```

```

Query Match          100.0%; Score 21; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
OY 1 SYDA 4
    ||||
DB 1 SYDA 4
```

```

RESULT 34
PCT-US02-11643-4149
; Sequence 4149, Application PC/TUS0211643
; GENERAL INFORMATION:
; APPLICANT: Challita-Eld, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Rubert, Rene S.
; APPLICANT: Morrison, Karen J.
; APPLICANT: Morrison, Robert K.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20067.40
; CURRENT APPLICATION NUMBER: PCT/US02/11643
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60,286,630
; NUMBER OF SEQ ID NOS: 15017
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4149
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapiens
PCT-US02-11643-4149

```

```

Query Match          100.0%; Score 21; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
OY 1 SYDA 4
    ||||
DB 2 SYDA 5
```

```

RESULT 35
PCT-US02-11643-4204
; Sequence 4204, Application PC/TUS0211643
; GENERAL INFORMATION:
; APPLICANT: Challita-Eld, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Rubert, Rene S.
; APPLICANT: Morrison, Karen J.
; APPLICANT: Morrison, Robert K.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20067.40
; CURRENT APPLICATION NUMBER: PCT/US02/11643
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60,286,630
; NUMBER OF SEQ ID NOS: 15017
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4204
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapiens
PCT-US02-11643-4204

```

```

Query Match          100.0%; Score 21; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
OY 1 SYDA 4
    ||||
```

Db 3 SYDA 6

RESULT 36  
PCT-US02-11643-4385; Sequence 4385, Application PC/TUS0211643  
; GENERAL INFORMATION:

; APPLICANT: Challita-Eld, Pia M.

; APPLICANT: Raitano, Arthur B.

; APPLICANT: Faris, Mary

; APPLICANT: Rubert, Rene S.

; APPLICANT: Morrison, Karen J.

; APPLICANT: Morrison, Robert K.

; APPLICANT: Ge, Wangmao

; APPLICANT: Jakobovits, Aya

; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN

; FILE REFERENCE: 51158-20067.40

; CURRENT APPLICATION NUMBER: PCT/US02/11643

; CURRENT FILING DATE: 2002-04-09

; PRIOR APPLICATION NUMBER: US 60/282,739

; PRIOR FILING DATE: 2001-04-10

; PRIOR APPLICATION NUMBER: US 60,286,630

; PRIOR FILING DATE: 2001-04-25

; NUMBER OF SEQ ID NOS: 15017

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4385

; LENGTH: 9

; TYPE: PRT

; ORGANISM: Homo Sapiens

PCT-US02-11643-4385

Query Match 100.0%; Score 21; DB 1; Length 9;

Best Local Similarity 100.0%; Pred. No. 4.2e+06;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4

Db 1 SYDA 4

RESULT 37  
PCT-US02-11643-4448

; Sequence 4448, Application PC/TUS0211643

; GENERAL INFORMATION:

; APPLICANT: Challita-Eld, Pia M.

; APPLICANT: Raitano, Arthur B.

; APPLICANT: Faris, Mary

; APPLICANT: Rubert, Rene S.

; APPLICANT: Morrison, Karen J.

; APPLICANT: Morrison, Robert K.

; APPLICANT: Ge, Wangmao

; APPLICANT: Jakobovits, Aya

; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN

; FILE REFERENCE: 51158-20067.40

; CURRENT APPLICATION NUMBER: PCT/US02/11643

; CURRENT FILING DATE: 2002-04-09

; PRIOR APPLICATION NUMBER: US 60/282,739

; PRIOR FILING DATE: 2001-04-10

; PRIOR APPLICATION NUMBER: US 60,286,630

; PRIOR FILING DATE: 2001-04-25

; NUMBER OF SEQ ID NOS: 15017

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4448

; LENGTH: 9

; TYPE: PRT

; ORGANISM: Homo Sapiens

PCT-US02-11643-4448

Query Match 100.0%; Score 21; DB 1; Length 9;

Best Local Similarity 100.0%; Pred. No. 4.2e+06;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4

Db 3 SYDA 6

RESULT 38  
PCT-US02-11643-4532

; Sequence 4532, Application PC/TUS0211643

; GENERAL INFORMATION:

; APPLICANT: Challita-Eld, Pia M.

; APPLICANT: Raitano, Arthur B.

; APPLICANT: Faris, Mary

; APPLICANT: Rubert, Rene S.

; APPLICANT: Morrison, Karen J.

; APPLICANT: Morrison, Robert K.

; APPLICANT: Ge, Wangmao

; APPLICANT: Jakobovits, Aya

; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN

; FILE REFERENCE: 51158-20067.40

; CURRENT APPLICATION NUMBER: PCT/US02/11643

; CURRENT FILING DATE: 2002-04-09

; PRIOR APPLICATION NUMBER: US 60/282,739

; PRIOR FILING DATE: 2001-04-10

; PRIOR APPLICATION NUMBER: US 60,286,630

; PRIOR FILING DATE: 2001-04-25

; NUMBER OF SEQ ID NOS: 15017

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4532

; LENGTH: 9

; TYPE: PRT

; ORGANISM: Homo Sapiens

PCT-US02-11643-4532

Query Match 100.0%; Score 21; DB 1; Length 9;

Best Local Similarity 100.0%; Pred. No. 4.2e+06;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4

Db 6 SYDA 9

RESULT 39  
PCT-US02-11643-4647

; Sequence 4647, Application PC/TUS0211643

; GENERAL INFORMATION:

; APPLICANT: Challita-Eld, Pia M.

; APPLICANT: Raitano, Arthur B.

; APPLICANT: Faris, Mary

; APPLICANT: Rubert, Rene S.

; APPLICANT: Morrison, Karen J.

; APPLICANT: Morrison, Robert K.

; APPLICANT: Ge, Wangmao

; APPLICANT: Jakobovits, Aya

; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN

; FILE REFERENCE: 51158-20067.40

; CURRENT APPLICATION NUMBER: PCT/US02/11643

; CURRENT FILING DATE: 2002-04-09

; PRIOR APPLICATION NUMBER: US 60/282,739

; PRIOR FILING DATE: 2001-04-10

; PRIOR APPLICATION NUMBER: US 60,286,630

; PRIOR FILING DATE: 2001-04-25

; NUMBER OF SEQ ID NOS: 15017

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4647

; LENGTH: 9

; TYPE: PRT

; ORGANISM: Homo Sapiens

PCT-US02-11643-4647

Query Match 100.0%; Score 21; DB 1; Length 9;

Best Local Similarity 100.0%; Pred. No. 4.2e+06;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 100.0%; Score 21; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0;

OY 1 SYDA 4  
|||||  
Db 4 SYDA 7

RESULT 40  
PCT-US02-11643-4799

; Sequence 4799, Application PC/TUS0211643

; GENERAL INFORMATION:  
; APPLICANT: Chaillita-Eld, Pia M.  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Faris, Mary  
; APPLICANT: Rubert, Rene S.  
; APPLICANT: Morrison, Karen J.  
; APPLICANT: Morrison, Robert K.  
; APPLICANT: Ge, Wangmao  
; APPLICANT: Jakobovits, Aya  
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN  
; FILE REFERENCE: 51158-20067.40  
; CURRENT APPLICATION NUMBER: PCT/US02/11643  
; CURRENT FILING DATE: 2002-04-09  
; PRIOR APPLICATION NUMBER: US 60/282,739  
; PRIOR FILING DATE: 2001-04-10  
; PRIOR APPLICATION NUMBER: US 60,286,630  
; PRIOR FILING DATE: 2001-04-25  
; NUMBER OF SEQ ID NOS: 15017  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4799  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
PCT-US02-11643-4799

Query Match 100.0%; Score 21; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0;

OY 1 SYDA 4  
|||||  
Db 5 SYDA 8

RESULT 41  
PCT-US02-11643-4800

; Sequence 4800, Application PC/TUS0211643

; GENERAL INFORMATION:  
; APPLICANT: Chaillita-Eld, Pia M.  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Faris, Mary  
; APPLICANT: Rubert, Rene S.  
; APPLICANT: Morrison, Karen J.  
; APPLICANT: Morrison, Robert K.  
; APPLICANT: Ge, Wangmao  
; APPLICANT: Jakobovits, Aya  
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN  
; FILE REFERENCE: 51158-20067.40  
; CURRENT APPLICATION NUMBER: PCT/US02/11643  
; CURRENT FILING DATE: 2002-04-09  
; PRIOR APPLICATION NUMBER: US 60/282,739  
; PRIOR FILING DATE: 2001-04-10  
; PRIOR APPLICATION NUMBER: US 60,286,630  
; PRIOR FILING DATE: 2001-04-25  
; NUMBER OF SEQ ID NOS: 15017  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4800  
; LENGTH: 9  
; TYPE: PRT

; ORGANISM: Homo Sapiens  
PCT-US02-11643-4800

Query Match 100.0%; Score 21; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0;

OY 1 SYDA 4  
|||||  
Db 2 SYDA 5

RESULT 42  
PCT-US02-11643-4949

; Sequence 4949, Application PC/TUS0211643  
; GENERAL INFORMATION:  
; APPLICANT: Chaillita-Eld, Pia M.  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Faris, Mary  
; APPLICANT: Rubert, Rene S.  
; APPLICANT: Morrison, Karen J.  
; APPLICANT: Morrison, Robert K.  
; APPLICANT: Ge, Wangmao  
; APPLICANT: Jakobovits, Aya  
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN  
; FILE REFERENCE: 51158-20067.40  
; CURRENT APPLICATION NUMBER: PCT/US02/11643  
; CURRENT FILING DATE: 2002-04-09  
; PRIOR APPLICATION NUMBER: US 60/282,739  
; PRIOR FILING DATE: 2001-04-10  
; PRIOR APPLICATION NUMBER: US 60,286,630  
; PRIOR FILING DATE: 2001-04-25  
; NUMBER OF SEQ ID NOS: 15017  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4949  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
PCT-US02-11643-4949

Query Match 100.0%; Score 21; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0;

OY 1 SYDA 4  
|||||  
Db 1 SYDA 4

RESULT 43  
PCT-US02-11643-5109

; Sequence 5109, Application PC/TUS0211643  
; GENERAL INFORMATION:  
; APPLICANT: Chaillita-Eld, Pia M.  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Faris, Mary  
; APPLICANT: Rubert, Rene S.  
; APPLICANT: Morrison, Karen J.  
; APPLICANT: Morrison, Robert K.  
; APPLICANT: Ge, Wangmao  
; APPLICANT: Jakobovits, Aya  
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN  
; FILE REFERENCE: 51158-20067.40  
; CURRENT APPLICATION NUMBER: PCT/US02/11643  
; CURRENT FILING DATE: 2002-04-09  
; PRIOR APPLICATION NUMBER: US 60/282,739  
; PRIOR FILING DATE: 2001-04-10  
; PRIOR APPLICATION NUMBER: US 60,286,630  
; PRIOR FILING DATE: 2001-04-25  
; NUMBER OF SEQ ID NOS: 15017  
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5109  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
PCT-US02-11643-5109

Query Match  
Best Local Similarity 100.0%; Score 21; DB 1; Length 9;  
Pred. No. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
||||  
DB 6 SYDA 9

RESULT 44  
PCT-US02-11643-5244

; Sequence 5244, Application PC/TUS0211643  
; GENERAL INFORMATION:  
; APPLICANT: Chaillita-Eid, Pia M.  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Faris, Mary  
; APPLICANT: Rubert, Rene S.  
; APPLICANT: Morrison, Karen J.  
; APPLICANT: Morrison, Robert K.  
; APPLICANT: Ge, Wangmao  
; APPLICANT: Jakobovits, Aya  
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN  
; FILE REFERENCE: 51158-20067.40  
; CURRENT APPLICATION NUMBER: PCT/US02/11643  
; CURRENT FILING DATE: 2002-04-09  
; PRIOR APPLICATION NUMBER: US 60/282,739  
; PRIOR FILING DATE: 2001-04-10  
; PRIOR APPLICATION NUMBER: US 60,286,630  
; PRIOR FILING DATE: 2001-04-25  
; NUMBER OF SEQ ID NOS: 15017  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5244  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
PCT-US02-11643-5244

Query Match  
Best Local Similarity 100.0%; Score 21; DB 1; Length 9;  
Pred. No. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
||||  
DB 2 SYDA 5

RESULT 45  
PCT-US02-11643-5290

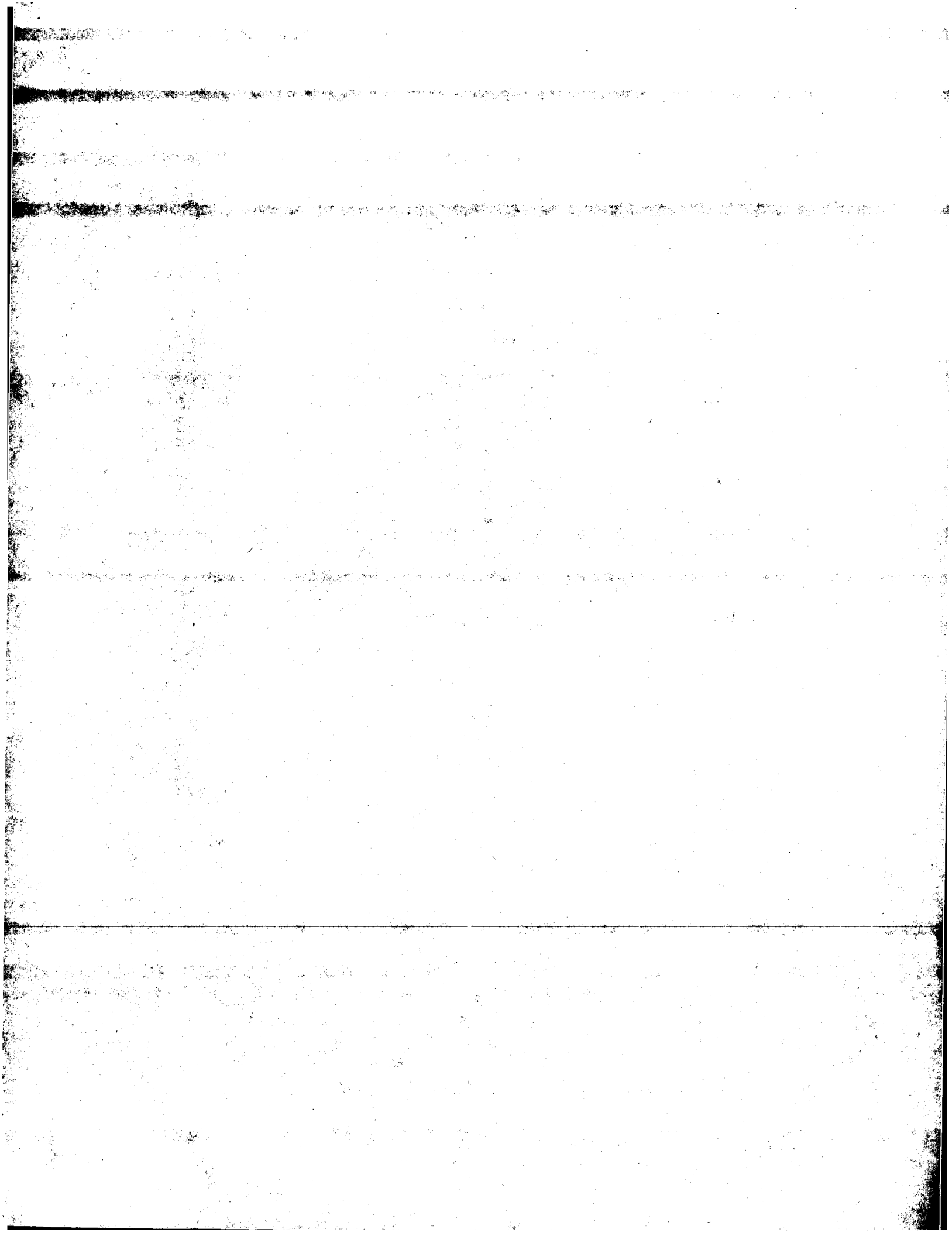
; Sequence 5290, Application PC/TUS0211643  
; GENERAL INFORMATION:  
; APPLICANT: Chaillita-Eid, Pia M.  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Faris, Mary  
; APPLICANT: Rubert, Rene S.  
; APPLICANT: Morrison, Karen J.  
; APPLICANT: Morrison, Robert K.  
; APPLICANT: Ge, Wangmao  
; APPLICANT: Jakobovits, Aya  
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN  
; FILE REFERENCE: 51158-20067.40  
; CURRENT APPLICATION NUMBER: PCT/US02/11643  
; CURRENT FILING DATE: 2002-04-09  
; PRIOR APPLICATION NUMBER: US 60/282,739  
; PRIOR FILING DATE: 2001-04-10  
; PRIOR APPLICATION NUMBER: US 60,286,630

; PRIOR FILING DATE: 2001-04-25  
; NUMBER OF SEQ ID NOS: 15017  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5290  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
PCT-US02-11643-5290

Query Match  
Best Local Similarity 100.0%; Score 21; DB 1; Length 9;  
Pred. No. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
||||  
DB 3 SYDA 6

Search completed: February 6, 2003, 11:37:33  
Job time : 134.167 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: February 6, 2003, 11:19:49 : Search time 14 Seconds  
(without alignments)  
23.165 Million cell updates/sec

Title: PAT943-3  
Perfect score: 21  
Sequence: 1 syda 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 405691 seqs, 81078759 residues

Total number of hits satisfying chosen parameters: 422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%  
Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	100.0	64	1 PCT-US02-41613-292	Sequence 292, App
2	21	100.0	65	6 US-10-203-138A-11128	Sequence 11128, A
3	21	100.0	65	6 US-10-203-138A-11208	Sequence 11208, A
4	21	100.0	65	6 US-10-276-781-1979	Sequence 1979, Ap
5	21	100.0	73	5 US-09-724-676-67729	Sequence 67729, A
6	21	100.0	73	5 US-09-724-676A-67729	Sequence 67729, A
7	21	100.0	75	5 US-09-513-999C-6382	Sequence 6382, Ap
8	21	100.0	77	5 US-09-513-999C-6627	Sequence 6627, Ap
9	21	100.0	98	5 US-09-724-676-67724	Sequence 67724, A
10	21	100.0	98	5 US-09-724-676-67725	Sequence 67725, A
11	21	100.0	98	5 US-09-724-676-67726	Sequence 67726, A
12	21	100.0	98	5 US-09-724-676-67728	Sequence 67728, A
13	21	100.0	98	5 US-09-724-676A-67724	Sequence 67724, A
14	21	100.0	98	5 US-09-724-676A-67725	Sequence 67725, A
15	21	100.0	98	5 US-09-724-676A-67726	Sequence 67726, A
16	21	100.0	98	5 US-09-724-676A-67728	Sequence 67728, A
17	21	100.0	127	1 PCT-US02-32727-6955	Sequence 6955, Ap
18	21	100.0	127	6 US-10-057-498-9695	Sequence 9695, Ap
19	21	100.0	128	1 PCT-US02-32727-25509	Sequence 25509, A
20	21	100.0	128	6 US-10-057-498-25509	Sequence 25509, A
21	21	100.0	130	5 US-09-950-084-7306	Sequence 7306, Ap
22	21	100.0	157	1 PCT-US02-32727-29763	Sequence 29763, A
23	21	100.0	159	5 US-09-724-676-67734	Sequence 67734, A
24	21	100.0	159	5 US-09-724-676A-67734	Sequence 67734, A
25	21	100.0	166	1 PCT-US02-32727-24114	Sequence 24114, A
26	21	100.0	166	1 PCT-US02-24310-27	Sequence 27, Appl

27	21	100.0	166	6 US-10-057-498-24114	Sequence 24114, A
28	21	100.0	177	5 US-09-134-000C-6245	Sequence 6245, Ap
29	21	100.0	177	5 US-09-134-000C-6245	Sequence 6245, Ap
30	21	100.0	184	5 US-09-724-676-67730	Sequence 67730, A
31	21	100.0	184	5 US-09-724-676-67731	Sequence 67731, A
32	21	100.0	184	5 US-09-724-676-67732	Sequence 67732, A
33	21	100.0	184	5 US-09-724-676-67733	Sequence 67733, A
34	21	100.0	184	5 US-09-724-676A-67730	Sequence 67730, A
35	21	100.0	184	5 US-09-724-676A-67731	Sequence 67731, A
36	21	100.0	184	5 US-09-724-676A-67732	Sequence 67732, A
37	21	100.0	184	5 US-09-724-676A-67733	Sequence 67733, A
38	21	100.0	199	6 US-10-218-140-2328	Sequence 2328, Ap
39	21	100.0	206	1 PCT-US02-32727-16492	Sequence 16492, A
40	21	100.0	206	6 US-10-057-498-16492	Sequence 16492, A
41	21	100.0	213	5 US-09-134-000C-3952	Sequence 3952, Ap
42	21	100.0	213	5 US-09-134-000C-3952	Sequence 3952, Ap
43	21	100.0	224	5 US-09-724-676-97153	Sequence 97153, A
44	21	100.0	224	5 US-09-724-676A-97153	Sequence 97153, A
45	21	100.0	231	5 US-09-724-676-97142	Sequence 97142, A

## ALIGNMENTS

RESULT 1  
PCT-US02-41613-292  
Sequence 292, Application PC/TUS0241613  
GENERAL INFORMATION:  
APPLICANT: diadexus, Inc.  
APPLICANT: Macina, Roberto  
APPLICANT: Sun, Yongming  
APPLICANT: Liu, Chenghua  
TITLE OF INVENTION: Compositions and Methods Relating to Urinary Specific Genes an  
FILE REFERENCE: DEX-0370  
CURRENT APPLICATION NUMBER: PCT/US02/41613  
PRIOR FILING DATE: 2002-12-23  
PRIOR APPLICATION NUMBER: 60/343,690  
PRIOR FILING DATE: 2001-12-28  
NUMBER OF SEQ ID NOS: 388  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO: 292  
LENGTH: 64  
TYPE: PRT  
ORGANISM: Homo sapien  
PCT-US02-41613-292

Query Match 100.0%; Score 21; DB 1; Length 64;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
DB 18 SYDA 21

RESULT 2  
US-10-203-138A-11128  
Sequence 11128, Application US/10203138A  
GENERAL INFORMATION:  
APPLICANT: Molecular Dynamics, Inc.  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO  
FILE REFERENCE: PB 0004 WO 8  
CURRENT APPLICATION NUMBER: US/10/203,138A  
CURRENT FILING DATE: 2002-08-02  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 04 February 2000 (04.02.00)  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 26 May 2000 (26.05.00)

```
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 03 August 2000 (03.08.00)
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 03 October 2000 (03.10.00)
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 27 September 2000 (27.09.00)
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 21 September 2000 (21.09.00)
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 30 June 2000 (30.06.00)
;; NUMBER OF SEQ ID NOS: 15438
;; SOFTWARE: Molecular Dynamics Sequence Listing Engine
;; SEQ ID NO 11128
;; LENGTH: 65
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AP000188.1
;; FEATURE:
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2
;; FEATURE:
;; OTHER INFORMATION: SWISSPROT HIT: P22415, EVALU 4.00e-14
;; FEATURE:
;; OTHER INFORMATION: EST_HUMAN HIT: AW937895.1, EVALU 1.00e-13
US-10-203-138A-11128
```

```
Query Match
Best Local Similarity 100.0%; Score 21; DB 6; Length 65;
Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 SYDA 4
    ||||
Db 28 SYDA 31
```

```
RESULT 3
US-10-203-138A-11208
; Sequence 11208, Application US/10203138A
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: PB 0004 WO 8
; CURRENT APPLICATION NUMBER: US/10/203,138A
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 15438
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
;; SEQ ID NO 11208
;; LENGTH: 65
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AP000044.1
;; FEATURE:
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.7
```

```
;; FEATURE:
;; OTHER INFORMATION: EST_HUMAN HIT: AW937895.1, EVALU 1.00e-16
;; FEATURE:
;; OTHER INFORMATION: SWISSPROT HIT: P22415, EVALU 3.00e-17
US-10-203-138A-11208
```

```
Query Match
Best Local Similarity 100.0%; Score 21; DB 6; Length 65;
Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 SYDA 4
    ||||
Db 36 SYDA 39
```

```
RESULT 4
US-10-276-781-1979
; Sequence 1979, Application US/10276781
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang et al.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-018 (785 contig)
; CURRENT APPLICATION NUMBER: US/10/276,781
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; NUMBER OF SEQ ID NOS: 2018
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1979
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-276-781-1979
```

```
Query Match
Best Local Similarity 100.0%; Score 21; DB 6; Length 65;
Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 SYDA 4
    ||||
Db 37 SYDA 40
```

```
RESULT 5
US-09-724-676-67729
; Sequence 67729, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 67729
;; LENGTH: 73
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc-feature
;; LOCATION: (11)..(12)
;; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
;; NAME/KEY: misc-feature
;; LOCATION: (14)..(14)
;; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-09-724-676-67729
```

```
Query Match
Best Local Similarity 100.0%; Score 21; DB 5; Length 73;
Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

OY 1 SYDA 4  
||||  
Db 54 SYDA 57

## RESULT 6

US-09-724-676A-67729  
; Sequence 67729, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724, 676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 67729  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (11)..(12)  
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: (14)..(14)  
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
US-09-724-676A-67729

Query Match  
Best Local Similarity 100.0%; Score 21; DB 5; Length 73;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
||||  
Db 54 SYDA 57

## RESULT 7

US-09-513-999C-6382  
; Sequence 6382, Application US/09513999C  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclet, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513, 999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 6382  
; LENGTH: 75  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-513-999C-6382

Query Match  
Best Local Similarity 100.0%; Score 21; DB 5; Length 75;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
||||  
Db 64 SYDA 67

RESULT 8  
US-09-513-999C-6627  
; Sequence 6627, Application US/09513999C  
; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Duclet, A.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

; FILE REFERENCE: 59.US2.REG

; CURRENT APPLICATION NUMBER: US/09/513, 999C

; CURRENT FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/122,487

; PRIOR FILING DATE: 1999-02-26

; NUMBER OF SEQ ID NOS: 36681

; SOFTWARE: Patent.pm

; SEQ ID NO 6627

; LENGTH: 77

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-513-999C-6627

Query Match  
Best Local Similarity 100.0%; Score 21; DB 5; Length 77;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
||||  
Db 7 SYDA 10

## RESULT 9

US-09-724-676-67724  
; Sequence 67724, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724, 676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 67724  
; LENGTH: 98  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (11)..(12)  
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: (14)..(14)  
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
US-09-724-676-67724

Query Match  
Best Local Similarity 100.0%; Score 21; DB 5; Length 98;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
||||  
Db 54 SYDA 57

## RESULT 10

US-09-724-676-67725  
; Sequence 67725, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724, 676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 67725  
; LENGTH: 98

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11)..(12)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14)..(14)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-09-724-676-67725
```

```
Query Match
Best Local Similarity 100.0%; Score 21; DB 5; Length 98;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYDA 4
    ||||
Db 54 SYDA 57
```

```

RESULT 11
US-09-724-676-67726
; Sequence 67726, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 67726
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11)..(12)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14)..(14)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-09-724-676-67726
```

```
Query Match
Best Local Similarity 100.0%; Score 21; DB 5; Length 98;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYDA 4
    ||||
Db 54 SYDA 57
```

```

RESULT 12
US-09-724-676-67728
; Sequence 67728, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 67728
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11)..(12)
```

```

; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14)..(14)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-09-724-676-67728
```

```
Query Match
Best Local Similarity 100.0%; Score 21; DB 5; Length 98;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYDA 4
    ||||
Db 54 SYDA 57
```

```

RESULT 13
US-09-724-676A-67724
; Sequence 67724, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 67724
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11)..(12)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14)..(14)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-09-724-676A-67724
```

```
Query Match
Best Local Similarity 100.0%; Score 21; DB 5; Length 98;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYDA 4
    ||||
Db 54 SYDA 57
```

```

RESULT 14
US-09-724-676A-67725
; Sequence 67725, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 67725
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11)..(12)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14)..(14)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
```

US-09-724-676A-67725

Query Match 100.0%; Score 21; DB 5; Length 98;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
||||  
DB 54 SYDA 57

RESULT 15

US-09-724-676A-67726  
; Sequence 67726, Application US/09724676A  
; GENERAL INFORMATION:

; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676A  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 67726  
; LENGTH: 98

; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (11)..(12)  
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (14)..(14)  
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
US-09-724-676A-67726

Query Match 100.0%; Score 21; DB 5; Length 98;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
||||  
DB 54 SYDA 57

RESULT 16

US-09-724-676A-67728  
; Sequence 67728, Application US/09724676A  
; GENERAL INFORMATION:

; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 67728  
; LENGTH: 98

; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (11)..(12)  
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (14)..(14)  
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
US-09-724-676A-67728

Query Match 100.0%; Score 21; DB 5; Length 98;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
||||  
DB 54 SYDA 57

RESULT 17

PCT-US02-32727-9695  
; Sequence 9695, Application PC/TUS0232727  
; GENERAL INFORMATION:

; APPLICANT: Mitcham, Jennifer  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Persing, David  
; APPLICANT: Bhatia, Ajay  
; APPLICANT: Maisonneuve, Jean Francois  
; APPLICANT: Zhang, Yanni  
; APPLICANT: Wang, Siqing  
; APPLICANT: Jen, Shyian  
; APPLICANT: Lodges, Michael  
; APPLICANT: Benson, Darin  
; APPLICANT: Jones, Robert  
; APPLICANT: Carter, Darlick  
; APPLICANT: Barth, Brenda  
; APPLICANT: Douglas, John  
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne  
; FILE REFERENCE: 210121.514C1  
; CURRENT APPLICATION NUMBER: PCT/US02/32727  
; CURRENT FILING DATE: 2002-10-11  
; NUMBER OF SEQ ID NOS: 30992  
; SEQ ID NO 9695  
; LENGTH: 127  
; TYPE: PRT  
; ORGANISM: Propionibacterium  
PCT-US02-32727-9695

Query Match 100.0%; Score 21; DB 1; Length 127;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
||||  
DB 106 SYDA 109

RESULT 18

US-10-057-498-9695  
; Sequence 9695, Application US/10057498  
; GENERAL INFORMATION:

; APPLICANT: Mitcham, Jennifer  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Persing, David  
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne  
; FILE REFERENCE: 210121.514  
; CURRENT APPLICATION NUMBER: US/10/057,498  
; CURRENT FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 29212  
; SEQ ID NO 9695  
; LENGTH: 127  
; TYPE: PRT  
; ORGANISM: Propionibacterium  
US-10-057-498-9695

Query Match 100.0%; Score 21; DB 6; Length 127;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
||||  
DB 106 SYDA 109

RESULT 19  
PCT-US02-32727-25509

```
; Sequence 25509, Application PC/TUS0232727
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siqing
; APPLICANT: Jen, Shyian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darlick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 25509
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Propionibacterium acnes
PCT-US02-32727-25509
```

```
Query Match          100.0%; Score 21; DB 1; Length 128;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYDA 4
      ||||
Db 49 SYDA 52
```

```
RESULT 20
US-10-057-498-25509
; Sequence 25509, Application US/10057498
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes
; FILE REFERENCE: 210121.514
; CURRENT APPLICATION NUMBER: US/10/057,498
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 29212
; SEQ ID NO 25509
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Propionibacterium acnes
US-10-057-498-25509
```

```
Query Match          100.0%; Score 21; DB 6; Length 128;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYDA 4
      ||||
Db 49 SYDA 52
```

```
RESULT 21
US-09-950-084-7306
; Sequence 7306, Application US/09950084
; GENERAL INFORMATION:
; APPLICANT: George H. Shimer, Jr.
; APPLICANT: George H. Miller
; APPLICANT: Roberta S. Hare
; APPLICANT: Karen J. Shaw
; TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods
; FILE REFERENCE: 1034/IC9630S2
```

```
; CURRENT APPLICATION NUMBER: US/09/950,084
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US 09/417,811
; PRIOR FILING DATE: 1999-10-14
; PRIOR APPLICATION NUMBER: US 09/353,718
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/266,557
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,556
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,555
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,542
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,541
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/037,934
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: US 09/036,720
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,338
; PRIOR FILING DATE: 1998-03-06
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 7451
; SEQ ID NO 7306
; LENGTH: 130
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-950-084-7306
```

```
Query Match          100.0%; Score 21; DB 5; Length 130;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYDA 4
      ||||
Db 67 SYDA 70
```

```
RESULT 22
PCT-US02-32727-29763
; Sequence 29763, Application PC/TUS0232727
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siqing
; APPLICANT: Jen, Shyian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darlick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 29763
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Propionibacterium acnes
PCT-US02-32727-29763
```

```
Query Match          100.0%; Score 21; DB 1; Length 157;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYDA 4
```

Db 100 SYDA 103

RESULT 23  
US-09-724-676-67734

; Sequence 67734, Application US/09724676

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 67734

; LENGTH: 159

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-724-676-67734

Query Match 100.0%; Score 21; DB 5; Length 159;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
1111  
Db 140 SYDA 143

RESULT 24  
US-09-724-676A-67734

; Sequence 67734, Application US/09724676A

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676A

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 67734

; LENGTH: 159

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-724-676A-67734

Query Match 100.0%; Score 21; DB 5; Length 159;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
1111  
Db 140 SYDA 143

RESULT 25  
PCT-US02-32727-24114

; Sequence 24114, Application PC/TUS032727

; GENERAL INFORMATION:

; APPLICANT: Mitcham, Jennifer

; APPLICANT: Skeiky, Yasir

; APPLICANT: Persing, David

; APPLICANT: Bhalla, Ajay

; APPLICANT: Maisonneuve, Jean Francois

; APPLICANT: Zhang, Yanni

; APPLICANT: Wang, Siqing

; APPLICANT: Jen, Shylan

; APPLICANT: Lodes, Michael

; APPLICANT: Benson, Darin

; APPLICANT: Jones, Robert

; APPLICANT: Carter, Darlick

; APPLICANT: Barth, Brenda

; APPLICANT: Douglass, John

; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne

; FILE REFERENCE: 210121.514c1

; CURRENT APPLICATION NUMBER: PCT/US02/32727

; CURRENT FILING DATE: 2002-10-11

; NUMBER OF SEQ ID NOS: 30992

; SEQ ID NO 24114

; LENGTH: 166

; TYPE: PRT

; ORGANISM: Propionibacterium

PCT-US02-32727-24114

Query Match 100.0%; Score 21; DB 1; Length 166;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
1111  
Db 118 SYDA 121

RESULT 26  
PCT-US02-24310-27

; Sequence 27, Application PC/TUS0224310

; GENERAL INFORMATION:

; APPLICANT: CUBIST PHARMACEUTICALS, INC.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS RELATING TO THE DAPTOMYCIN

; FILE REFERENCE: CUB-12 PCT CIP

; CURRENT APPLICATION NUMBER: PCT/US02/24310

; CURRENT FILING DATE: 2002-10-25

; PRIOR APPLICATION NUMBER: PCT/US01/32354

; PRIOR FILING DATE: 2001-10-17

; PRIOR APPLICATION NUMBER: 60/310,385

; PRIOR FILING DATE: 2001-08-06

; PRIOR APPLICATION NUMBER: 60/379,866

; PRIOR FILING DATE: 2002-05-10

; NUMBER OF SEQ ID NOS: 170

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 27

; LENGTH: 166

; TYPE: PRT

; ORGANISM: Streptomyces roseosporus

PCT-US02-24310-27

Query Match 100.0%; Score 21; DB 1; Length 166;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
1111  
Db 66 SYDA 69

RESULT 27  
US-10-057-498-24114

; Sequence 24114, Application US/10057498

; GENERAL INFORMATION:

; APPLICANT: Mitcham, Jennifer

; APPLICANT: Skeiky, Yasir

; APPLICANT: Persing, David

; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne

; FILE REFERENCE: 210121.514

; CURRENT APPLICATION NUMBER: US/10/057,498

; CURRENT FILING DATE: 2001-04-20

; NUMBER OF SEQ ID NOS: 29212

; SEQ ID NO 24114

; LENGTH: 166

; TYPE: PRT

; ORGANISM: Propionibacterium

US-10-057-498-24114

Query Match 100.0%; Score 21; DB 6; Length 166;

Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
||||  
Db 118 SYDA 121

RESULT 28  
US-09-134-000C-6245

; Sequence 6245, Application US/09134000C  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; FILE REFERENCE: 032796-032  
; CURRENT APPLICATION NUMBER: US/09/134,000C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055,778  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6245  
; LENGTH: 177  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-134-000C-6245

Query Match 100.0%; Score 21; DB 5; Length 177;  
Best Local Similarity 100.0%; Pred. No. 4.5e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
||||  
Db 155 SYDA 158

RESULT 29

US-09-134-000C-6245  
; Sequence 6245, Application US/09134000C  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; FILE REFERENCE: 032796-032  
; CURRENT APPLICATION NUMBER: US/09/134,000C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055,778  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6245  
; LENGTH: 177  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-134-000C-6245

Query Match 100.0%; Score 21; DB 5; Length 177;  
Best Local Similarity 100.0%; Pred. No. 4.5e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
||||  
Db 155 SYDA 158

RESULT 30  
US-09-724-676-67730

; Sequence 67730, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 67730  
; LENGTH: 184  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676-67730

Query Match 100.0%; Score 21; DB 5; Length 184;  
Best Local Similarity 100.0%; Pred. No. 4.6e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
||||  
Db 140 SYDA 143

RESULT 31  
US-09-724-676-67731

; Sequence 67731, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 67731  
; LENGTH: 184  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676-67731

Query Match 100.0%; Score 21; DB 5; Length 184;  
Best Local Similarity 100.0%; Pred. No. 4.6e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
||||  
Db 140 SYDA 143

RESULT 32

US-09-724-676-67732  
; Sequence 67732, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 67732  
; LENGTH: 184  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676-67732

Query Match 100.0%; Score 21; DB 5; Length 184;  
Best Local Similarity 100.0%; Pred. No. 4.6e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
||||  
Db 140 SYDA 143

RESULT 33

US-09-724-676-67733



```

; Sequence 67733, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 67733
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-67733

```

```

Query Match          100.0%; Score 21; DB 5; Length 184;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      1 SYDA 4
      ||||
Db      140 SYDA 143

```

```

RESULT 34
US-09-724-676A-67730
; Sequence 67730, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 67730
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-67730

```

```

Query Match          100.0%; Score 21; DB 5; Length 184;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      1 SYDA 4
      ||||
Db      140 SYDA 143

```

```

RESULT 35
US-09-724-676A-67731
; Sequence 67731, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 67731
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-67731

```

```

Query Match          100.0%; Score 21; DB 5; Length 184;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      1 SYDA 4
      ||||

```

```

Db      140 SYDA 143

```

```

RESULT 36
US-09-724-676A-67732
; Sequence 67732, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 67732
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-67732

```

```

Query Match          100.0%; Score 21; DB 5; Length 184;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      1 SYDA 4
      ||||
Db      140 SYDA 143

```

```

RESULT 37
US-09-724-676A-67733
; Sequence 67733, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 67733
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-67733

```

```

Query Match          100.0%; Score 21; DB 5; Length 184;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      1 SYDA 4
      ||||
Db      140 SYDA 143

```

```

RESULT 38
US-10-218-140-2328
; Sequence 2328, Application US/10218140
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES
; FILE REFERENCE: 15966-543 CON
; CURRENT APPLICATION NUMBER: US/10/218,140
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 09/540,763
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/127,728
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 60/127,636
; PRIOR FILING DATE: 1999-04-02
; PRIOR APPLICATION NUMBER: 60/127,607

```

;; PRIOR FILING DATE: 1999-03-31  
;; NUMBER OF SEQ ID NOS: 6322  
;; SOFTWARE: Curanator Version 1.0  
;; SEQ ID NO 2328  
;; LENGTH: 199  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-218-140-2328

Query Match  
Best Local Similarity 100.0%; Score 21; DB 6; Length 199;  
Pred. No. 5.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
|||  
DB 8 SYDA 11

RESULT 39

PCT-US02-32727-16492  
; Sequence 16492, Application PC/TUS0232727  
; GENERAL INFORMATION:  
; APPLICANT: Mitcham, Jennifer  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Persing, David  
; APPLICANT: Bhatia, Ajay  
; APPLICANT: Maisonneuve, Jean Francois  
; APPLICANT: Zhang, Yanni  
; APPLICANT: Wang, Sigang  
; APPLICANT: Jen, Shylan  
; APPLICANT: Lodes, Michael  
; APPLICANT: Benson, Darin  
; APPLICANT: Jones, Robert  
; APPLICANT: Carter, Darlick  
; APPLICANT: Barth, Brenda  
; APPLICANT: Douglas, John  
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes  
; FILE REFERENCE: 210121.514c1  
; CURRENT APPLICATION NUMBER: PCT/US02/32727  
; PRIOR FILING DATE: 2002-10-11  
; NUMBER OF SEQ ID NOS: 30992  
; SEQ ID NO 16492  
; LENGTH: 206  
; TYPE: PRT  
; ORGANISM: Propionibacterium acnes  
PCT-US02-32727-16492

Query Match  
Best Local Similarity 100.0%; Score 21; DB 1; Length 206;  
Pred. No. 5.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
|||  
DB 97 SYDA 100

RESULT 40

US-10-057-498-16492  
; Sequence 16492, Application US/10057498  
; GENERAL INFORMATION:  
; APPLICANT: Mitcham, Jennifer  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Persing, David  
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes  
; FILE REFERENCE: 210121.514  
; CURRENT APPLICATION NUMBER: US/10/057,498  
; PRIOR FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 29212  
; SEQ ID NO 16492  
; LENGTH: 206  
; TYPE: PRT  
; ORGANISM: Propionibacterium acnes  
US-10-057-498-16492

Query Match  
Best Local Similarity 100.0%; Score 21; DB 6; Length 206;  
Pred. No. 5.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
|||  
DB 97 SYDA 100

RESULT 41

US-09-134-000C-3952  
; Sequence 3952, Application US/09134000C  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; FILE REFERENCE: 032796-032  
; CURRENT APPLICATION NUMBER: US/09/134,000C  
; PRIOR FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055,778  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3952  
; LENGTH: 213  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-134-000C-3952

Query Match  
Best Local Similarity 100.0%; Score 21; DB 5; Length 213;  
Pred. No. 5.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
|||  
DB 157 SYDA 160

RESULT 42

US-09-134-000C-3952  
; Sequence 3952, Application US/09134000C  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; FILE REFERENCE: 032796-032  
; CURRENT APPLICATION NUMBER: US/09/134,000C  
; PRIOR FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055,778  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3952  
; LENGTH: 213  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-134-000C-3952

Query Match  
Best Local Similarity 100.0%; Score 21; DB 5; Length 213;  
Pred. No. 5.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
|||  
DB 157 SYDA 160

RESULT 43

US-09-724-676-97153  
; Sequence 97153, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD

Search completed: February 6, 2003, 11:39:02  
job time : 15 secs

```

; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 97153
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-97153
    
```

```

Query Match
Best Local Similarity 100.0%; Score 21; DB 5; Length 224;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    
```

```

OY 1 SYDA 4
    ||||
DB 189 SYDA 192
    
```

```

RESULT 44
US-09-724-676A-97153
; Sequence 97153, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 97153
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-97153
    
```

```

Query Match
Best Local Similarity 100.0%; Score 21; DB 5; Length 224;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    
```

```

OY 1 SYDA 4
    ||||
DB 189 SYDA 192
    
```

```

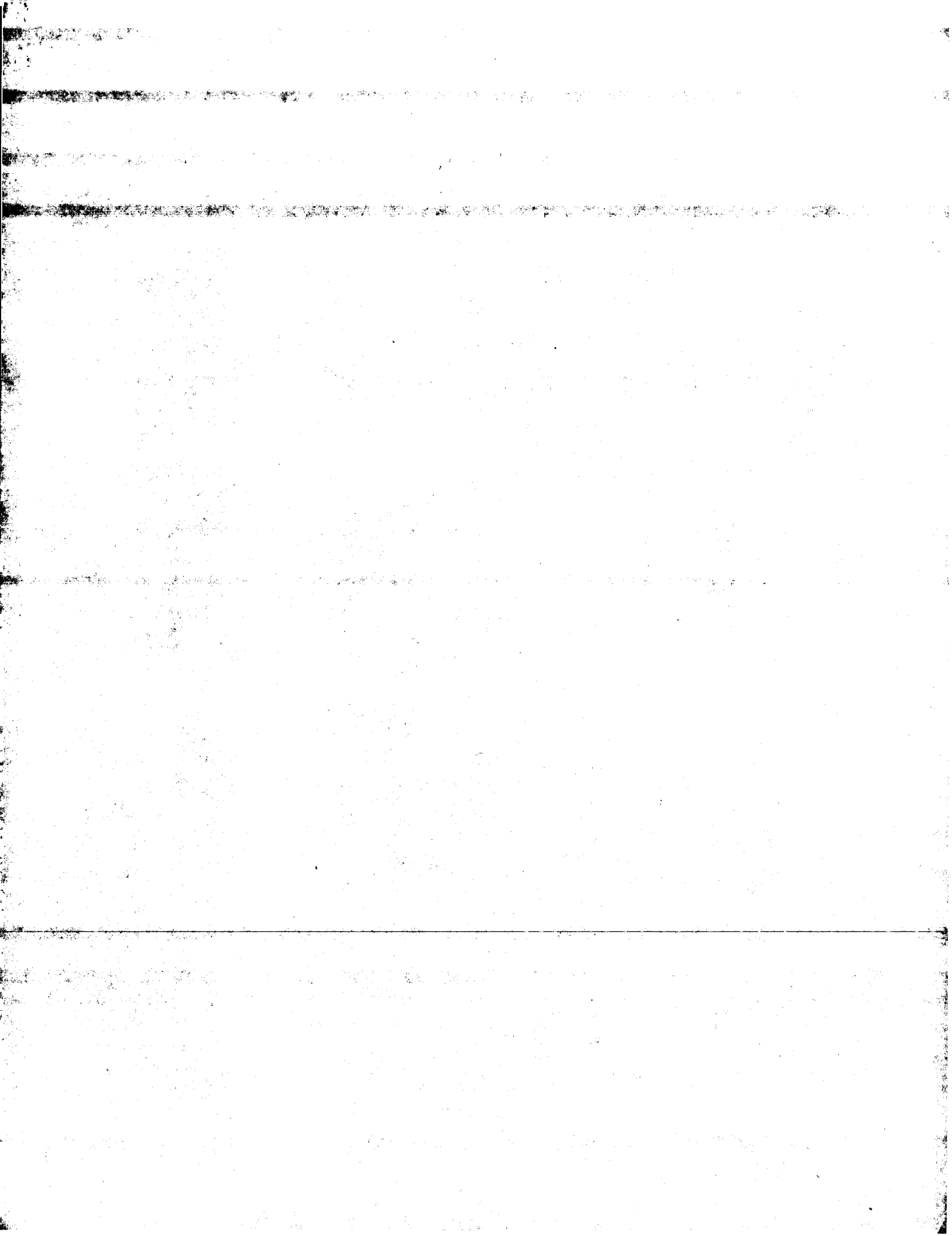
RESULT 45
US-09-724-676-97142
; Sequence 97142, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 97142
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-97142
    
```

```

Query Match
Best Local Similarity 100.0%; Score 21; DB 5; Length 231;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    
```

```

OY 1 SYDA 4
    ||||
DB 196 SYDA 199
    
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 11:16:04 ; Search time 10.5 Seconds  
(without alignments)  
36.623 Million cell updates/sec

Title: PAT943-3  
Perfect score: 21  
Sequence: 1 syda 4

Scoring table: BL0SUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 667

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%  
Maximum Match 100%  
Listing first 45 summaries

Database :

PIR\_73:\*  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	15	2	PH1582
2	21	100.0	53	2	AB1896
3	21	100.0	67	2	A59104
4	21	100.0	80	2	AC1788
5	21	100.0	81	2	TL4992
6	21	100.0	81	2	AD1412
7	21	100.0	84	2	P90808
8	21	100.0	84	2	B85568
9	21	100.0	84	2	A64849
10	21	100.0	85	2	AB0638
11	21	100.0	85	2	JS0644
12	21	100.0	98	2	T31168
13	21	100.0	97	2	C47673
14	21	100.0	103	2	T31207
15	21	100.0	106	2	H84370
16	21	100.0	113	2	A75626
17	21	100.0	113	2	D97729
18	21	100.0	121	2	H95968
19	21	100.0	127	2	T23550
20	21	100.0	128	2	C90733
21	21	100.0	129	2	S77545
22	21	100.0	130	2	C89791
23	21	100.0	132	2	E85583
24	21	100.0	133	2	G97136
25	21	100.0	133	2	T35218
26	21	100.0	133	2	T26268
27	21	100.0	133	2	D97240
28	21	100.0	134	2	B87162
29	21	100.0	135	2	I46643

#### ALIGNMENTS

```

30      21      100.0      139      2      E97426      flagellar basal-bo
31      21      100.0      139      2      AE2644      flagellar basal-bo
32      21      100.0      139      2      B64010      hypothetical prote
33      21      100.0      139      2      F70661      probable lipoprote
34      21      100.0      143      1      A44397      coflin - yeast (S
35      21      100.0      144      2      T44916      superoxide dismuta
36      21      100.0      144      2      H97356      rearranged T-cell
37      21      100.0      144      2      H97356      single strand DNA
38      21      100.0      144      2      A64662      ATP synthase F0, s
39      21      100.0      145      2      A71856      ATP synthase b' -
40      21      100.0      145      2      H89587      protein R09F10.5 (
41      21      100.0      145      2      F81272      hypothetical prote
42      21      100.0      148      2      E95886      probable sugar-pho
43      21      100.0      158      2      T22645      hypothetical prote
44      21      100.0      159      2      G70939      probable hsp prote
45      21      100.0      159      2      G97723      hypothetical prote

```

#### RESULT 1

```

PH1582
Ig H Chain V-D-J region (wild-type clone 5) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Jun-1994 #sequence_rev150 02-Jun-1994 #text_change 17-Mar-1999
C:Accession: PH1582
R:Levinson, D.A.; Campos-Torres, J.; Leder, P.
J. Exp. Med. 178, 317-329, 1993
A:Title: Molecular characterization of transgene-induced immunodeficiency in B-1ess m
A:Reference number: PH1580; MUID:93301609; PMID:8315387
A:Accession: PH1582
A:Molecule type: DNA
A:Residues: 1-15 <LEV>
A:Experimental source: bone marrow pre-B lymphocyte
C:Keywords: Immunoglobulin

```

Query Match 100.0%; Score 21; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
||||  
DB 9 SYDA 12

#### RESULT 2

```

AB1896
hypothetical protein asl0716 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_rev150 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AB1896
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kurihara, T.; Sasamoto, S.; Watanabe, A.; Irigun
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A:Reference number: AB1897; MUID:21595285; PMID:11759840
A:Accession: AB1896
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-53 <KUR>
A:Cross-references: GB:BA000019; PIDN:BAH72673.1; PID:g17130061; GSPDB:GN00179
A:Experimental source: strain PCC 7120
A:Genetics: asl0716

```

Query Match 100.0%; Score 21; DB 2; Length 53;  
Best Local Similarity 100.0%; Pred. No. 98;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
||||

Db 2 SYDA 5

RESULT 3

A59104

hypothetical protein pX01-105 - *Bacillus anthracis* virulence plasmid pX01  
C:Species: *Bacillus anthracis*  
C:Date: 12-Nov-1999 #sequence\_revision 12-Nov-1999 #text\_change 11-May-2000  
C:Accession: A59104

R:Okimaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Kelm, P.; Koehler

J. Bacteriol. 181, 6509-6515, 1999

A:Title: Sequence and organization of pX01, the large *Bacillus anthracis* plasmid harbored

A:Reference number: A59091; MUID:9445483; PMID:10515943

A:Accession: A59104

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-67 <OK>

A:Cross-references: GB:AF065404; NID:94894216; PIDN:AMD32409.1; PID:94894321

A:Experimental source: strain Sterne

A:Note: similar to transcription state regulatory protein, abr, B. subtilus (P08874). OF

C:Genetics:

A:Gene: pX01-105

A:Genome: plasmid

Query Match 100.0%; Score 21; DB 2; Length 67;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4

Db 25 SYDA 28

RESULT 4

AC1788

B. subtilis Yaal protein homolog lin2849 [imported] - *Listeria innocua* (strain C11p11262

C:Species: *Listeria innocua*

C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001

C:Accession: AC1788

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkatt, G.; Madueno, E.; Maitournam, A.; Ma

ok, C.; Schluter, T.; Simoes, N.; Tiller, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,

A:Title: Comparative genomics of *Listeria* species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AC1788

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-80 <GLA>

A:Cross-references: GB:AL592022; PIDN:CAC98075.1; PID:916415384; GSPDB:GN00178

A:Experimental source: strain C11p11262

C:Genetics:

Query Match 100.0%; Score 21; DB 2; Length 80;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4

Db 22 SYDA 25

RESULT 5

T14992

hypothetical protein Y1075 - *Yersinia pestis* plasmid pMT1

C:Species: *Yersinia pestis*

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999

C:Accession: T14992

R:Rindler, L.E.; Plano, G.V.; Burland, V.; Mayhew, G.F.; Blattner, F.R.

Infect. Immun. 66, 5731-5742, 1998

A:Title: Complete DNA sequence and detailed analysis of the *Yersinia pestis* KIM5 plas

A:Reference number: Z18268; MUID:99043898; PMID:9826348

A:Accession: T14992

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-81 <LIN>

A:Cross-references: EMBL:AF074611; NID:93883003; PID:93883075; PIDN:AC082735.1

C:Genetics:

A:Gene: Y1075

A:Genome: plasmid pMT1

Query Match 100.0%; Score 21; DB 2; Length 81;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4

Db 42 SYDA 45

RESULT 6

AD1412

B. subtilis Yaal protein homolog lmo2701 [imported] - *Listeria monocytogenes* (strain

C:Species: *Listeria monocytogenes*

C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001

C:Accession: AD1412

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkatt, G.; Madueno, E.; Maitournam, A.; Ma

ok, C.; Schluter, T.; Simoes, N.; Tiller, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla

A:Title: Comparative genomics of *Listeria* species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AD1412

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-81 <GLA>

A:Cross-references: GB:NC\_003210; PIDN:CAD00914.1; PID:916412201; GSPDB:GN00177

A:Experimental source: strain ESD-e

C:Genetics:

A:Gene: lmo2701

Query Match 100.0%; Score 21; DB 2; Length 81;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4

Db 22 SYDA 25

RESULT 7

F90808

hypothetical protein Ecsl438 [imported] - *Escherichia coli* (strain O157:H7, substrain

C:Species: *Escherichia coli*

C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001

C:Accession: F90808

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C

gasawara, N.; Yasunaga, T.; Kunara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and g

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: F90808

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-84 <HAY>

A:Cross-references: GB:BA000007; PIDN:BA034861.1; PID:913360902; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain RMD 0509952

C:Genetics:

A:Gene: Ecsl438

Query Match 100.0%; Score 21; DB 2; Length 84;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
||||  
DB 23 SYDA 26

## RESULT 8

B85668

hypothetical protein ycep [imported] - Escherichia coli (strain O157:H7, substrain EDL93)

C:Species: Escherichia coli

C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001

C:Accession: B85668

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,

Nature 409, 529-533, 2001

A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: B85668

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-84 <STO>

A:Cross-references: GB:AE005174; NID:912514595; PIDN:AAG55806.1; GSPDB:GN00145; UWGP:216

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: ycep

Query Match 100.0%; Score 21; DB 2; Length 84;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
||||  
DB 23 SYDA 26

## RESULT 9

A64849

yeep protein - Escherichia coli (strain K-12)

C:Species: Escherichia coli

C>Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 01-Mar-2002

C:Accession: A64849

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A>Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: A64849

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-84 <BLAT>

A:Cross-references: GB:AE000207; GB:U00096; NID:91787293; PIDN:AACT4144.1; PID:91787299;

A:Experimental source: strain K-12, substrain MG1655

C:Genetics:

A:Gene: ycep

Query Match 100.0%; Score 21; DB 2; Length 84;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
||||  
DB 23 SYDA 26

## RESULT 10

AB0638

conserved hypothetical protein STY1199 [imported] - Salmonella enterica subsp. enterica

C:Species: Salmonella enterica subsp. enterica serovar typhi

A>Note: this species has also been called Salmonella typhi

C>Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 09-Nov-2001

C:Accession: AB0638

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church  
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr  
, S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,

A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se

A:Reference number: AB0502; PMID:11677608

A:Accession: AB0638

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-85 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD08285.1; PID:916502331; GSPDB:GN00176

C:Genetics:

A:Gene: STY1199

Query Match 100.0%; Score 21; DB 2; Length 85;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
||||  
DB 24 SYDA 27

## RESULT 11

J50644

eclosion hormone precursor - silkworm

C:Species: Bombyx mori (silkworm)

C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 21-Jul-2000

C:Accession: J50644

R:Kanito, T.; Tanaka, H.; Sato, B.; Nagasawa, H.; Suzuki, A.

Biochem. Biophys. Res. Commun. 182, 514-519, 1992

A>Title: Nucleotide sequence of cDNA for the eclosion hormone of the silkworm, Bombyx

A:Reference number: J50644; MUID:92134263; PMID:1370883

A:Accession: J50644

A:Molecule type: mRNA

A:Residues: 1-88 <KAM>

A:Cross-references: GB:D10135; NID:9217271; PIDN:BA01012.1; PID:9217272

C:Comment: This neuropeptide controls ecdysis.

C:Superfamily: eclosion hormone

F:1-26/Domain: signal sequence #status predicted <SIG>

F:27-88/Product: eclosion hormone #status predicted <ECL>

Query Match 100.0%; Score 21; DB 2; Length 88;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
||||  
DB 33 SYDA 36

## RESULT 12

T31168

hypothetical protein 392 - Sphingomonas aromaticivorans plasmid pNL1

C:Species: Sphingomonas aromaticivorans

C>Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 11-Jan-2000

C:Accession: T31168

R:Romine, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.

submitted to the EMBL Data Library, July 1998

A:Description: Complete sequence of a 184 kb catabolic plasmid from Sphingomonas arom

A:Reference number: Z20992

A:Accession: T31168

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-95 <ROM>

A:Cross-references: EMBL:AF079317; NID:9378261; PID:9378309; PIDN:AD03892.1

C:Genetics:

A:Genome: plasmid pNL1

A>Note: ori392

Query Match 100.0%; Score 21; DB 2; Length 95;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SYDA 4  
 ||||  
 Db 57 SYDA 60

## RESULT 13

C47673  
 ferredoxin [2Fe-2S] - *Synechococcus* sp. (PCC 7002)  
 C:Species: *Synechococcus* sp.  
 C:Date: 19-Dec-1993 #sequence\_revision 18-Nov-1994 #text\_change 11-Jun-1999  
 C:Accession: C47673  
 R:Leonhardt, K.; Straus, N.A.  
 J. Gen. Microbiol. 138, 1613-1621, 1992  
 A:Title: An iron stress operon involved in photosynthetic electron transport in the mar  
 A:Reference number: A47673; MUID:92407507; PMID:1527503  
 A:Accession: C47673  
 A:Status: preliminary  
 A:Molecule type: nucleic acid  
 A:Residues: 1-97 <DEO>  
 A:Cross-references: GB:M88251; NID:g154545; PIDN:AAA27329.1; PID:g154546  
 A:Experimental source: PCC 7002  
 A:Note: sequence extracted from NCBI backbone (NCBIN:113308, MCBIP:113309)  
 C:Superfamily: ferredoxin [2Fe-2S]; ferredoxin [2Fe-2S] homology  
 C:Keywords: 2Fe-2S; electron transfer; iron-sulfur protein; metalloprotein  
 F:25-79/Domain: ferredoxin [2Fe-2S] homology <FER>  
 F:40,45,48,78/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Query Match  
 Best Local Similarity 100.0%; Score 21; DB 2; Length 97;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
 ||||  
 Db 16 SYDA 19

## RESULT 14

T31207  
 hypothetical protein 680 - *Sphingomonas aromaticivorans* plasmid pNL1  
 C:Species: *Sphingomonas aromaticivorans*  
 C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 18-Feb-2000  
 C:Accession: T31207  
 R:Romine, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.; G  
 submitted to the EMBL Data Library, July 1998  
 A:Description: Complete sequence of a 184 kb catabolic plasmid from *Sphingomonas aromati*  
 A:Reference number: Z20992  
 A:Accession: T31207  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-103 <ROM>  
 A:Cross-references: EMBL:AF079317; NID:g3378261; PID:g3378348; PIDN:AA03931.1  
 C:Genetics:  
 A:Genome: plasmid pNL1  
 A:Note: ori680  
 C:Superfamily: *Sphingomonas aromaticivorans* plasmid pNL1 hypothetical protein 680

Query Match  
 Best Local Similarity 100.0%; Score 21; DB 2; Length 103;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
 ||||  
 Db 17 SYDA 20

## RESULT 15

H84370  
 hypothetical protein Vng2202h [imported] - *Halobacterium* sp. NRC-1  
 C:Species: *Halobacterium* sp. NRC-1  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C:Accession: H84370

R.Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky  
 ; Letthauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Ja  
 Jung, K.H.; Alam, M.; Freitas, T.  
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.,  
 A:Title: Genome sequence of *Halobacterium* species NRC-1.  
 A:Reference number: A84160; MUID:20504483; PMID:11016950  
 A:Accession: H84370  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-106 <SNO>  
 A:Cross-references: GB:AE004437; NID:g10581619; PIDN:AA020332.1; GSPDB:GN00138  
 C:Genetics:  
 A:Gene: VNG2202H

Query Match  
 Best Local Similarity 100.0%; Score 21; DB 2; Length 106;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
 ||||  
 Db 78 SYDA 81

## RESULT 16

A75626  
 salicylate monooxygenase-related protein - *Deinococcus radiodurans* (strain R1)  
 C:Species: *Deinococcus radiodurans*  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
 C:Accession: A75626  
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J  
 ; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.  
 A:Reference number: A75250; MUID:20036896; PMID:10567266  
 A:Accession: A75626  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1113 <WHI>  
 A:Cross-references: GB:AE001826; NID:g6460827; PIDN:AAF12594.1; PID:g6460890; TIGR:DR  
 A:Experimental source: strain R1  
 C:Genetics:  
 A:Gene: DRB0072  
 A:Map position: megaplasmid  
 A:Genome: plasmid  
 A:Note: plasmid MP1

Query Match  
 Best Local Similarity 100.0%; Score 21; DB 2; Length 113;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
 ||||  
 Db 104 SYDA 107

## RESULT 17

D97729  
 hypothetical protein RC0236 [imported] - *Rickettsia conorii* (strain Malish 7)  
 C:Species: *Rickettsia conorii*  
 C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 30-Sep-2001  
 C:Accession: D97729  
 R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barde, V.; Samson, D.,  
 Science 293, 2093-2098, 2001  
 A:Title: Mechanisms of Evolution in *Rickettsia conorii* and *Rickettsia prowazekii*.  
 A:Reference number: A97700; MUID:21442074; PMID:11557893  
 A:Accession: D97729  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-115 <KUR>  
 A:Cross-references: GB:AE006914; PIDN:AA02774.1; PID:g15619290; GSPDB:GN00173  
 C:Genetics:



A:Gene: RC0236

Query Match 100.0%; Score 21; DB 2; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 DB 87 SYDA 90

## RESULT 18

H95968

probable regulatory protein, atsr family [imported] - *Shinorhizobium meliloti* (strain 102)  
 C:Species: *Shinorhizobium meliloti*  
 C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001

C:Accession: H95968  
 R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chalm, P.; Vorholter, F.J.; Hernan  
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
 A:Title: The complete sequence of the 1,683-kb PSymb megaplasmid from the N2-fixing endo  
 A:Reference number: A95842; MUID:2136508; PMID:11481431

A:Accession: H95968  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-121 <RUR>

A:Cross-references: GB:AL591985; PIDN:CAC49416.1; PID:q15140902; GSPDB:GN00167  
 A:Experimental source: strain 1021, megaplasmid PSymb

R:Galbert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
 pella, D.; Chalm, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
 L.; Hyman, R.W.; Jones, T.  
 Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
 hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Mong, K.; Yeh, K.  
 A:Title: The composite genome of the legume symbiont *Shinorhizobium meliloti*.  
 A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation  
 C:Genetics:  
 A:Gene: Smb21576  
 A:Genome: plasmid

Query Match 100.0%; Score 21; DB 2; Length 121;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 DB 76 SYDA 79

## RESULT 19

T23550

hypothetical protein K09G1.2 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Mar-2000

C:Accession: T23550  
 R:Lloyd, C.  
 submitted to the EMBL Data Library, October 1996

A:Reference number: Z19758  
 A:Accession: T23550  
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
 A:Residues: 1-127 <NTL>

A:Cross-references: EMBL:261101; PIDN:CAR03198.1; GSPDB:GN00023; CESP:K09G1.2  
 A:Experimental source: clone K09G1  
 C:Genetics:  
 A:Gene: CESP:K09G1.2  
 A:Map position: 5  
 A:Introns: 25/3

C:Superfamily: *Caenorhabditis elegans* hypothetical protein K09G1.2

Query Match 100.0%; Score 21; DB 2; Length 127;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 DB 88 SYDA 91

## RESULT 20

C90733

probable minor tail protein [imported] - *Escherichia coli* (strain O157:H7, substrain  
 C:Species: *Escherichia coli*  
 C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001

C:Accession: C90733  
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C  
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
 DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and g  
 A:Reference number: A9629; MUID:21156231; PMID:11258796  
 A:Accession: C90733  
 A:Status: preliminary

A:Molecule type: DNA  
 A:Residues: 1-128 <HAV>

A:Cross-references: GB:BA000007; PIDN:BAR34258.1; PID:q13360294; GSPDB:GN00154  
 A:Experimental source: strain O157:H7, substrain RMD 0509952  
 C:Genetics:  
 A:Gene: ECs0835

C:Superfamily: phage lambda minor tail protein C

Query Match 100.0%; Score 21; DB 2; Length 128;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 DB 92 SYDA 95

## RESULT 21

S7545

hypothetical protein slr1210 - *Synechocystis* sp. (strain PCC 6803)  
 C:Species: *Synechocystis* sp.  
 A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
 C:Accession: S7545  
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,  
 O. K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas  
 DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocys*  
 s.  
 A:Reference number: S74322; MUID:97061201; PMID:8905231  
 A:Accession: S7545  
 A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA  
 A:Residues: 1-129 <KAN>  
 A:Cross-references: EMBL:D90905; GB:AB001339; NID:q1652360; PIDN:BAI17392.1; PID:d101

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 100.0%; Score 21; DB 2; Length 129;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 DB 28 SYDA 31

## RESULT 22

C89791

hypothetical protein [imported] - *Staphylococcus aureus* (strain N315)

C:Species: *Staphylococcus aureus*  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001

C:Accession: C89791  
 R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O  
 ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
 Lancet 357, 1225-1240, 2001  
 A>Title: Whole genome sequencing of metacillin-resistant *Staphylococcus aureus*.  
 A:Reference number: A89758; MUID:21311952; PMID:11418146  
 A:Accession: C89791  
 A>Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-130 <KUR>  
 A:Cross-references: GB:BA000018; PID:g13700188; PIDN:BA841486.1; GSPDB:GN00149  
 A:Experimental source: strain N315  
 C:Genetics:  
 A:Gene: SA0262

Query Match 100.0%; Score 21; DB 2; Length 130;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 Db 67 SYDA 70

## RESULT 23

E85583

Probable tail component of prophage CP-933K 20973 [Imported] - *Escherichia coli* (strain  
 C:Species: *Escherichia coli*  
 C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
 A:Accession: E85583  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 Miller, L.; Grotbeck, E.J.; Davis, N.W.; Llim, A.; Dinalanta, E.; Potamousis, K.; Apodaca,  
 Nature 409, 529-533, 2001  
 A>Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A:Accession: E85583  
 A>Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-132 <STO>  
 A:Cross-references: GB:AE005174; NID:g12513741; PIDN:AA655129.1; GSPDB:GN00145; UWGP:Z09  
 A:Experimental source: strain O157:H7, substrain EDL933  
 C:Genetics:  
 A:Gene: Z0973  
 C:Superfamily: phage lambda minor tail protein G

Query Match 100.0%; Score 21; DB 2; Length 132;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 Db 96 SYDA 99

## RESULT 24

G97136

Phage related SSB-like protein [Imported] - *Clostridium acetobutylicum*  
 C:Species: *Clostridium acetobutylicum*  
 C>Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 17-May-2002  
 A:Accession: G97136  
 R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
 J. Bacteriol. 183, 4823-4838, 2001  
 A>Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium *Clo*  
 A:Reference number: A96900; MUID:21359325; PMID:21359325  
 A:Accession: G97136  
 A>Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-133 <KUR>  
 A:Cross-references: GB:AE001437; PIDN:AAK79882.1; PID:g15024899; GSPDB:GN00168  
 A:Experimental source: *Clostridium acetobutylicum* ATCC824  
 C:Genetics:  
 A:Gene: CAC319  
 C:Superfamily: bacterial single-stranded DNA-binding protein; single-stranded DNA-bindin

Query Match 100.0%; Score 21; DB 2; Length 133;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 Db 80 SYDA 83

## RESULT 25

T35218

hypothetical protein SC5C7.11 SC5C7.11 Streptomyces coelicolor  
 C:Species: *Streptomyces coelicolor*  
 C>Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
 A:Accession: T35218  
 R:Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
 submitted to the EMBL Data Library, September 1998  
 A:Reference number: Z21572  
 A:Accession: T35218  
 A>Status: Preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-133 <SEE>  
 A:Cross-references: EMBL:AL031515; PIDN:CAA20623.1; GSPDB:GN00070; SCOEDB:SC5C7.11  
 A:Experimental source: strain A3(2)  
 C:Genetics:  
 A:Gene: SCOEDB:SC5C7.11

Query Match 100.0%; Score 21; DB 2; Length 133;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 Db 110 SYDA 113

## RESULT 26

T26268

hypothetical protein W07G1.4 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 A:Accession: T26268  
 R:Percy, C.  
 submitted to the EMBL Data Library, November 1996  
 A:Reference number: Z20186  
 A:Accession: T26268  
 A>Status: Preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-133 <WIT>  
 A:Cross-references: EMBL:Z82076; NID:e1247330; PIDN:CAB04936.1; GSPDB:GN00020; CESP:W  
 A:Experimental source: clone W07G1  
 C:Genetics:  
 A:Gene: CESP:W07G1.4  
 A:Map position: 2  
 A:Insertions: 36/2; 60/3

Query Match 100.0%; Score 21; DB 2; Length 133;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 Db 87 SYDA 90

## RESULT 27

D97240

uncharacterized conserved protein CAC2767 [Imported] - *Clostridium acetobutylicum*  
 C:Species: *Clostridium acetobutylicum*  
 C>Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
 A:Accession: D97240  
 R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L  
 .; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001  
 A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld  
 A:Reference number: A96900; MUID:21359325; PMID:21359325  
 A:Accession: D97240  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-133 <RUR>  
 A:Cross-references: GB:AE001437; PIDN:AAK80711.1; PID:q15025804; GSPDB:GN00168  
 A:Experimental source: Clostridium acetobutylicum ATCC824  
 A:Gene: CAC2767

Query Match  
 Best Local Similarity 100.0%; Score 21; DB 2; Length 133;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 Db 56 SYDA 59

RESULT 28  
 B87162  
 Probable secreted protein [imported] - Mycobacterium leprae  
 C:Species: Mycobacterium leprae  
 C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
 C:Accession: B87162  
 R:Coile, S.T.; Eljilmet, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho  
 R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,  
 eam, M.A.; Rutherford, K.M.  
 Nature 409, 1007-1011, 2001  
 A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq  
 A:Title: Massive gene decay in the leprosy bacillus  
 A:Reference number: A86909; MUID:21128732; PMID:11234002  
 A:Accession: B87162  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-134 <STO>  
 A:Cross-references: GB:AL450380; NID:q13093649; PIDN:CAC30978.1; GSPDB:GN00147  
 A:Gene: ML2023

Query Match  
 Best Local Similarity 100.0%; Score 21; DB 2; Length 134;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 Db 36 SYDA 39

RESULT 29  
 I46643  
 rearranged T-cell receptor delta-chain/Vdelta5.2-Delta5-jdelta1 - pig (fragment)  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 21-Jan-2000  
 C:Accession: I46643  
 R:Tang, Y.G.; Ohta, S.; Yamada, S.; Shimizu, M.; Takagaki, Y.  
 J. Immunol. 155, 1981-1993, 1995  
 A:Title: Diversity of T cell receptor delta-chain cDNA in the thymus of a one-month-old  
 A:Reference number: I46623; MUID:95363165; PMID:7636249  
 A:Accession: I46643  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-135 <YAN>  
 A:Cross-references: GB:D49584; NID:q1041170; PIDN:BA08528.1; PID:q1041171  
 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology  
 C:Keywords: T-cell receptor  
 F:32-112/Domain: Immunoglobulin homology <IMM>

Query Match  
 Best Local Similarity 100.0%; Score 21; DB 2; Length 135;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 Db 72 SYDA 75

RESULT 30  
 E97426  
 flagellar basal-body rod protein flgC [imported] - Agrobacterium tumefaciens (strain  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 11-Jan-2002  
 C:Accession: E97426  
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm  
 A.; Liu, F.; Mollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz,  
 Science 294, 2323-2328, 2001  
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium  
 A:Reference number: A97359; PMID:11743194  
 A:Accession: E97426  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-139 <KUR>  
 A:Cross-references: GB:AE007869; PIDN:AAK86366.1; PID:q15155492; GSPDB:GN00169  
 A:Gene: AGR\_C\_975  
 A:Map position: circular chromosome  
 C:Superfamily: rod protein flgC

Query Match  
 Best Local Similarity 100.0%; Score 21; DB 2; Length 139;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 Db 113 SYDA 116

RESULT 31  
 AE2644  
 flagellar basal-body rod protein [imported] - Agrobacterium tumefaciens (strain C58,  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 01-Feb-2002  
 C:Accession: AE2644  
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Moo  
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McCl  
 ; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam  
 ster, E.W.  
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A:Reference number: AB2577; PMID:11743193  
 A:Accession: AE2644  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-139 <KUR>  
 A:Cross-references: GB:AE008688; PIDN:AL41571.1; PID:q17738905; GSPDB:GN00186  
 A:Experimental source: strain C58 (Dupont)  
 C:Genetics:  
 A:Gene: flgC  
 A:Map position: circular chromosome  
 C:Superfamily: rod protein flgC

Query Match  
 Best Local Similarity 100.0%; Score 21; DB 2; Length 139;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 Db 113 SYDA 116

RESULT 32  
 B64010  
 hypothetical protein HI0589 - Haemophilus influenzae (strain Rd KW20)

```
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 20-Sep-1999
C:Accession: B64010
R:Feilschman, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirnness, E.F.; Kerlavage, A.;
Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodde, A.; Kelley, J.M.; Weidman, J.;
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fulhmann, J.L.; Geoghagen, N.S.M.
Science 269, 446-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: B64010
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1139 <TIGR>
A:Cross-references: GB:U32740; GB:I42023; NID:91573572; PIDN:AAC22246.1; PID:91573579; T
C:Genetics:
A:Start codon: GTG
C:Superfamily: Escherichia coli sigma-E factor regulatory protein rsec

Query Match          100.0%; Score 21; DB 2; Length 139;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4
    ||||
Db 3 SYDA 6

RESULT 33
F70661
Probable lipoprotein - Mycobacterium tuberculosis (strain H37Rv)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: F70661
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
Connor, R.; Davies, R.; Devlin, R.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: F70661
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1139 <COL>
A:Cross-references: GB:283860; GB:AL123456; NID:93261681; PIDN:CAB06158.1; PID:e290748;
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: lppO

Query Match          100.0%; Score 21; DB 2; Length 139;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4
    ||||
Db 82 SYDA 85

RESULT 34
A44397
Coflin - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein L0595; protein YIL050C
C:Species: Saccharomyces cerevisiae
C:Date: 31-Dec-1993 #sequence_revision 27-Jun-1994 #text_change 21-Jul-2000
C:Accession: A44397; B44397; JN0529; S64802; S50970; S31309; S36087
R:Moore, A.L.; Janney, P.A.; Louie, K.A.; Drubin, D.G.
J. Cell Biol. 120, 421-435, 1993
A:Title: Coflin is an essential component of the yeast cortical cytoskeleton.
A:Reference number: A44397; MUID:93132073; PMID:8421056
A:Molecule type: DNA
A:Residues: 1-143 <MOO>

A:Cross-references: EMBL:Z14971; NID:93563; PIDN:CAA78694.1; PID:93564
A:Note: sequence extracted from NCBI backbone (NCBIN:122683, NCBIPI:122684)
A:Accession: B44397
A:Molecule type: protein
A:Residues: 43-56;83-96,'X',98;106-129,'DS',132-141 <MO2>
A:Note: sequence extracted from NCBI backbone
R:Ida, K.; Moriyama, K.; Matsumoto, S.; Kawasaki, H.; Nishida, E.; Yahara, I.
Gene 124, 115-120, 1993
A:Title: Isolation of a yeast essential gene, COF1, that encodes a homologue of mamma
A:Reference number: JN0529; MUID:93178959; PMID:8440472
A:Accession: JN0529
A:Molecule type: DNA
A:Residues: 1-143 <IID>
A:Cross-references: GB:D13230; NID:9287599; PIDN:BA02514.1; PID:9287600
R:Wedler, H.; Wedler, E.; Scharfe, M.; Wambutt, R.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64792
A:Accession: S64802
A:Molecule type: DNA
A:Residues: 1-143 <MED>
A:Cross-references: EMBL:Z73155; NID:91360250; PIDN:CAA97502.1; PID:91360251; GSPDB:G
R:Wedler, H.; Wambutt, R.
submitted to the EMBL Data Library, January 1995
A:Reference number: S50950
A:Accession: S50970
A:Molecule type: DNA
A:Residues: 6-143 <WED>
A:Cross-references: EMBL:Z47973; NID:9642313; PIDN:CAA88007.1; PID:9642334
C:Comment: Coflin reversibly regulates actin polymerization and depolymerization in
C:Genetics:
A:Gene: SGD:COF1; MIPS:YIL050C
A:Cross-references: SGD:S0003973; MIPS:YIL050C
A:Map position: 12L
A:Introns: 5/2
C:Superfamily: coflin
C:Keywords: actin binding
F:88-118/Region: actin binding #status predicted

Query Match          100.0%; Score 21; DB 1; Length 143;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4
    ||||
Db 49 SYDA 52

RESULT 35
T44916
Superoxide dismutase (EC 1.15.1.1) (Fe) [similarity] - Haloferax mediterranei (fragme
C:Species: Haloferax mediterranei
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 24-Oct-2000
C:Accession: T44916
R:Dennis, P.P.; Shimmitt, L.C.
Microbiol. Mol. Biol. Rev. 61, 90-104, 1997
A:Title: Evolutionary divergence and salinity-mediated selection in halophilic archae
A:Reference number: Z22865; MUID:97257269; PMID:9106366
A:Accession: T44916
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-144 <DEN>
A:Cross-references: EMBL:U078908; NID:92190597; PIDN:AB060932.1; PID:92190598
C:Genetics:
A:Gene: sod1
C:Superfamily: superoxide dismutase (Mn)
C:Keywords: metalloprotein; oxidoreductase

Query Match          100.0%; Score 21; DB 2; Length 144;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Oy 1 SYDA 4  
 ||||  
 Db 71 SYDA 74

# RESULT 36

146642  
 rearranged T-cell receptor delta-chain/Vdelta5.1-Ddelta5-udelta1 - pig (fragment)  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 23-Jul-1999  
 C:Accession: 146642  
 R:Yang, Y.G.; Ohta, S.; Yamada, S.; Shimizu, M.; Takagaki, Y.  
 J. Immunol. 155, 1981-1993, 1995  
 A:Title: Diversity of T cell receptor delta-chain cDNA in the thymus of a one-month-old  
 A:Reference number: 146623; MUID:95363165; PMID:7636249  
 A:Accession: 146642  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-144 <YAN>  
 A:Cross-references: GB:ID49583; NID:g1041168; PIDN:BA005527.1; PID:g1041169  
 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology  
 C:Keywords: T-cell receptor  
 F:32-112/Domain: Immunoglobulin homology <IMV>

Query Match 100.0%; Score 21; DB 2; Length 144;  
 Best Local Similarity 100.0%; Pred. No. 3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SYDA 4  
 ||||  
 Db 72 SYDA 75

# RESULT 37

H97356  
 single strand DNA binding protein, SSR [imported] - Clostridium acetobutylicum  
 C:Species: Clostridium acetobutylicum  
 C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 30-Sep-2001  
 C:Accession: H97356  
 R:Noelling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Kocain, E.V.; Smith, D.R.  
 J. Bacteriol. 183, 4823-4838, 2001  
 A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld  
 A:Reference number: A96900; MUID:21359325; PMID:21359325  
 A:Accession: H97356  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-144 <KUR>  
 A:Cross-references: GB:AE001437; PIDN:AAK81643.1; PID:g15026829; GSPDB:GN00168  
 A:Experimental source: Clostridium acetobutylicum ATCC824  
 C:Genetics: CAC3723  
 C:Superfamily: Bacterial single-stranded DNA-binding protein; single-stranded DNA-binding

Query Match 100.0%; Score 21; DB 2; Length 144;  
 Best Local Similarity 100.0%; Pred. No. 3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SYDA 4  
 ||||  
 Db 80 SYDA 83

# RESULT 38

A64662  
 ATP synthase F0, subunit b' - Helicobacter pylori (strain 26695)  
 C:Species: Helicobacter pylori  
 C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
 C:Accession: A64662  
 R:Tombs, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliq, H.G.; Glöck, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Matthey, L.; Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.; Fraser, A.; Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
 A:Reference number: A64520; MUID:97394467; PMID:9252185  
 A:Accession: A64662  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-144 <TOM>  
 A:Cross-references: GB:AE000619; GB:AE000511; NID:g2314276; PIDN:AAD08179.1; PID:g231

Query Match 100.0%; Score 21; DB 2; Length 144;  
 Best Local Similarity 100.0%; Pred. No. 3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SYDA 4  
 ||||  
 Db 89 SYDA 92

# RESULT 39

A71856  
 ATP synthase b' - Helicobacter pylori (strain J99)  
 C:Species: Helicobacter pylori  
 A:Variety: strain J99  
 C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 08-Oct-1999  
 C:Accession: A71856  
 R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.  
 Nature 397, 176-180, 1999  
 A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric p  
 A:Reference number: A71800; MUID:99120557; PMID:9923682  
 A:Accession: A71856  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-144 <ARN>  
 A:Cross-references: GB:AE001533; GB:AE001439; NID:g4155636; PIDN:AAD06630.1; PID:g415  
 A:Experimental source: strain J99  
 C:Genetics: A:Gene: atpx

Query Match 100.0%; Score 21; DB 2; Length 144;  
 Best Local Similarity 100.0%; Pred. No. 3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SYDA 4  
 ||||  
 Db 89 SYDA 92

# RESULT 40

H89587  
 Protein R09F10.5 [imported] - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Nov-2001  
 C:Accession: H89587  
 R:Anonymous, The C. elegans Sequencing Consortium.  
 Science 282, 2012-2018, 1998  
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio  
 A:Reference number: A75000; MUID:9905613; PMID:9851916  
 A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C-  
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;  
 A:Accession: H89587  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-145 <STO>  
 A:Cross-references: GB:chr\_X; PIDN:AAAG9094.1; PID:g1465853; GSPDB:GN00028; CESP:R09F  
 C:Genetics: A:Gene: R09F10.5  
 A:Map position: X

Query Match 100.0%; Score 21; DB 2; Length 145;  
 Best Local Similarity 100.0%; Pred. No. 3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
|||||  
Db 93 SYDA 96

## RESULT 41

F81272

hypothetical protein Cj1298 truncated homolog Cj1297 [similarity] - Campylobacter jejuni  
C:Species: Campylobacter jejuni  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 03-Jun-2002  
C:Accession: F81272  
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelley, J.M.; Churcher, C.; Basham, D.; Chilling  
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell  
Nature 403, 665-668, 2000  
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyf  
A:Reference number: A81250; MUID:20150912; PMID:10688204  
A:Accession: F81272  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-145 <PAR>  
A:Cross-references: GB:AL139078; GB:AL111168; NID:g6968723; PIDN:CAB73724.1; PID:g696873  
A:Experimental source: serotype O2, strain NCTC 11168  
C:Genetics:  
A:Gene: Cj1297

Query Match 100.0%; Score 21; DB 2; Length 145;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
|||||  
Db 118 SYDA 121

## RESULT 42

E95886

probable sugar-phosphate isomerase protein [imported] - Sinorhizobium meliloti (strain 1  
C:Species: Sinorhizobium meliloti  
C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
C:Accession: E95886  
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chalm, P.; Vorholter, F.J.; Hernan  
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
A:Title: The complete sequence of the 1,683-kb PSymb megaplasmid from the N2-fixing endo  
A:Reference number: A95842; MUID:21396508; PMID:11481431  
A:Accession: E95886  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-148 <KUR>  
A:Cross-references: GB:AL591985; PIDN:CAC48757.1; PID:g515140230; GSPDB:GN00167  
R:Experimental source: strain 1021, megaplasmid pSymb  
R:Gallbert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
P.; Chalm, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Leleure,  
hebaunt, P.; Vandemol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
A:Reference number: A96039; MUID:21368234; PMID:11474104  
A:Contents: annotation  
C:Genetics:  
A:Gene: Smb20371  
A:Superfamily: galactoside O-acetyltransferase

Query Match 100.0%; Score 21; DB 2; Length 148;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
|||||  
Db 34 SYDA 37

## RESULT 43

T22645

hypothetical protein F54D5.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T22645  
R:Coles, L.; Matthews, L.  
submitted to the EMBL Data Library, October 1995  
A:Reference number: Z19593  
A:Accession: T22645  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-158 <WIL>  
A:Cross-references: EMBL:Z66513; NID:e1062016; PIDN:CAA91330.1; GSPDB:GN00020; CESP:F  
A:Experimental source: clone F54D5  
C:Genetics:  
A:Gene: CESP:F54D5.3  
A:Map position: 2  
A:Introns: 70/1; 100/3; 132/3

Query Match 100.0%; Score 21; DB 2; Length 158;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
|||||  
Db 138 SYDA 141

## RESULT 44

G70939

probable hsp protein - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C:Accession: G70939  
R:Coole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon  
A:Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd,  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: G70939  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-159 <COL>  
A:Cross-references: GB:AL021929; GB:AL123456; NID:g3242291; PIDN:CAA17343.1; PID:g290  
A:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: hsp  
C:Superfamily: alpha-crystallin

Query Match 100.0%; Score 21; DB 2; Length 159;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
|||||  
Db 130 SYDA 133

## RESULT 45

G97723

hypothetical protein RC0191 [imported] - Rickettsia conorii (strain Malish 7)  
C:Species: Rickettsia conorii  
C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 22-Oct-2001  
C:Accession: G97723  
R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.;  
Science 293, 2093-2098, 2001  
A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.  
A:Reference number: A97700; MUID:21442074; PMID:11557893  
A:Accession: G97723  
A:Status: preliminary

A:Molecule type: DNA  
 A:Residues: 1-159 <RUR>  
 A:Cross-references: GB:AE006914; PIDN:AAU02729.1; PID:g15619240; GSPDB:GN00173  
 C:Genetics:  
 A:Gene: RC0191  
 C:Superfamily: 4-hydroxyphenylacetate 3-monooxygenase small chain  
  
 Query Match 100.0%; Score 21; DB 2; Length 159;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 1 SYDA 4  
 ||||  
 Db 124 SYDA 127  
  
 Search completed: February 6, 2003, 11:23:06  
 Job time : 12.5 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 11:16:00 : Search time 5.33333 Seconds

(without alignments)  
31.107 Million cell updates/sec

Title: PAT943-3  
Perfect score: 21  
Sequence: 1 syda 4

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 257

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 100%  
Maximum Match 100%  
Listing first 45 summaries

Database: SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	100.0	84	1 YCEP_ECOLI	P75927 escherichia
2	21	100.0	88	1 ECLH_BOMMO	P25331 bombyx mori
3	21	100.0	96	1 PERL_SYNP2	P31965 synchococc
4	21	100.0	109	1 PER_LOXAL	Q25221 loxocera al
5	21	100.0	139	1 FLGC_AGR5	Q44336 agrobacteri
6	21	100.0	139	1 Y589_HAEIN	P44020 haemophilus
7	21	100.0	143	1 COFI_YEAST	Q03408 saccharomyc
8	21	100.0	144	1 SOD1_HALME	O08461 halobacteri
9	21	100.0	153	1 GALX_CANMA	P56600 candida mal
10	21	100.0	156	1 RIB2_PHOE	O93692 photobacter
11	21	100.0	164	1 R12B_TRIBB	P41165 trypanosoma
12	21	100.0	167	1 HXD3_BRARE	O42370 brachydanio
13	21	100.0	168	1 SCYD_ASPEU	O14434 aspergillus
14	21	100.0	180	1 RBS2_MESCR	O04450 mesembryant
15	21	100.0	182	1 RBS1_MESCR	P16032 mesembryant
16	21	100.0	183	1 ATPD_BACND	O9K8h2 bacillus ha
17	21	100.0	183	1 RBS3_MESCR	Q08183 mesembryant
18	21	100.0	186	1 RBS6_MESCR	Q08186 mesembryant
19	21	100.0	202	1 SODF_NICPL	P22302 nicotiana p
20	21	100.0	205	1 FLID_PROMI	P42274 proteus mlt
21	21	100.0	208	1 FTSD_STRGR	P45503 streptomyce
22	21	100.0	210	1 HXB3_BRARE	O42368 brachydanio
23	21	100.0	216	1 ATKQ_HALNI	P57687 halobacteri
24	21	100.0	216	1 OL11_HUMAN	P47889 homo sapien
25	21	100.0	217	1 U119_CAEER	Q10658 caenorhabdi
26	21	100.0	219	1 U119_CAEER	Q10658 caenorhabdi
27	21	100.0	224	1 Y945_METJA	O58355 methanococc
28	21	100.0	227	1 YVVA_METTE	P29577 methanobact
29	21	100.0	230	1 YVIB_AGR9	P05357 agrobacteri
30	21	100.0	230	1 YVIB_AGR9	P05357 agrobacteri
31	21	100.0	236	1 Y416_CHLNR	O84421 chlamydia t
32	21	100.0	237	1 YVIB_AGR5	P17798 agrobacteri
33	21	100.0	250	1 PBCB_SCHGR	P82886 schistocerc

34	21	100.0	250	1 Y625_TREPA	O83633 treponema p
35	21	100.0	251	1 RL7C_SCHPO	O60143 schizosacch
36	21	100.0	253	1 SOL_TREPA	O83286 treponema p
37	21	100.0	255	1 BIOC_SERMA	P36571 seriatia ma
38	21	100.0	259	1 CAH2_CHICK	P07630 gallus gall
39	21	100.0	265	1 EXL2_ARATH	O9565 arabidopsis
40	21	100.0	266	1 FLIP_CAUCR	O45980 caulobacter
41	21	100.0	271	1 PAMB_CORGL	O9X732 corynebacte
42	21	100.0	274	1 CYNT_SYNY3	O54735 synchocyst
43	21	100.0	279	1 YS11_CAEEL	Q10033 caenorhabdi
44	21	100.0	280	1 SUHB_METTH	O26957 methanobact
45	21	100.0	281	1 DEGV_BACSV	P32436 bacillus su

## ALIGNMENTS

## RESULT 1

YCEP\_ECOLI STANDARD; PRT; 84 AA.

AC P75927; Q9R374; 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein ycep.

GN YCEP OR B1060.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OX Escherichia.

NCBI\_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=K12 / MG1655;

RX MEDLINE=97426617; PubMed=9278503;

RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,

RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,

RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

RA Mau B., Shao Y.,

RT "The complete genome sequence of Escherichia coli K-12.";

RL Science 277:1453-1474 (1997).

[2]

RP SEQUENCE FROM N.A.

RC STRAIN=K12;

RX MEDLINE=97061202; PubMed=8905232;

RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,

RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kishimoto K.,

RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,

RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,

RA Samped G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,

RA Yano M., Horiiuchi T.,

RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome

corresponding to the 12.7-28.0 min region on the linkage map.";

RL DNA Res. 3:137-155 (1996).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; AE000207; AAC7414.1; -

DR EMBL; D90742; BAA35857.1; ALT\_INIT.

DR EMBL; D90743; BAA35868.1; ALT\_INIT.

CC Ecogene: EGI1435; ycep.

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 84 AA; 9664 MW; 1D7103C964E5D5FF CRC64;

Query Match 100.0%; Score 21; DB 1; Length 84;

Best Local Similarity 100.0%; Pred. No. 65;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 DB 23 SYDA 26

## RESULT 2

ECLH\_BOMMO STANDARD; PRT; 88 AA.  
 ID ECLH\_BOMMO  
 AC P2531;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Ecdysion hormone precursor (Ecdysion activator) (EH).  
 OS Bombyx mori (Silk moth).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 OC Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;  
 OC Diptera; Bombycoidea; Bombycidae; Bombyx.  
 NCBI\_Taxid=7091;

SEQUENCE FROM N.A.  
 RX MEDLINE=92134263; PubMed=1370883;  
 RA Kamito T., Tanaka H., Sato B., Nagasawa H., Suzuki A.;  
 RT "Nucleotide sequence of cDNA for the ecdysion hormone of the  
 RT silkworm, Bombyx mori, and the expression in a brain."  
 RL Biochem. Biophys. Res. Commun. 182:514-519(1992).  
 RN [2]  
 RP SEQUENCE OF 27-87.  
 RA Kono T., Nagasawa H., Isogai A., Fugo H., Suzuki A.;  
 RT "Amino acid sequence of ecdysion hormone of the silkworm, Bombyx  
 RT mori.";  
 RL Agric. Biol. Chem. 51:2307-2308(1987).  
 CC -1- FUNCTION: NEUROPEPTIDE THAT TRIGGERS THE PERFORMANCE OF ECDYSIS  
 CC BEHAVIORS AT THE END OF A MOLT. IT TRIGGERS ADULT BEHAVIOR  
 CC PATTERNS: LARVAL, PUPAL AND ADULT ECDYSIS, AND PLASTICIZATION  
 CC DURING THE MOLT.  
 CC -1- SIMILARITY: HIGH. TO OTHER INSECTS ECDYSION HORMONE.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL; D10135; BAA01012.1; -;  
 DR PIR; JS0644; JS0644.  
 KW Hormone; Neuropeptide; Signal.

FT CHAIN 1 26  
 FT SIGNAL 1 26  
 FT DISULFID 27 88  
 FT DISULFID 40 64  
 FT DISULFID 44 60  
 FT DISULFID 47 75  
 FT CONFLICT 31 31  
 FT CONFLICT 33 33  
 FT CONFLICT 46 46  
 FT CONFLICT 67 68  
 FT CONFLICT 70 71  
 FT SEQUENCE 88 AA; 9505 MW; 429DDCOADPAD852 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 88;  
 Best Local Similarity 100.0%; Pred. No. 68;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 DB 33 SYDA 36

## RESULT 3

PERL\_SYNP2 STANDARD; PRT; 96 AA.  
 ID FERL\_SYNP2  
 AC P31965;

DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 01-JUL-1993 (Rel. 26, Last annotation update)  
 DE Ferredoxin I.  
 GN PERP.

OS Synechococcus sp. (Strain PCC 7002) (Agmenellum quadruplicatum).  
 OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.  
 NCBI\_Taxid=32049;

SEQUENCE FROM N.A.  
 RX MEDLINE=92407507; PubMed=1527503;  
 RA Leonhardt K.G., Straus N.A.;  
 RT "An iron stress operon involved in photosynthetic electron transport  
 RT in the marine cyanobacterium Synechococcus sp. PCC 7002.";  
 RL J. Gen. Microbiol. 138:1613-1621(1992).  
 CC -1- FUNCTION: FERREDOXINS ARE IRON-SULFUR PROTEINS THAT TRANSFER  
 CC ELECTRONS IN A WIDE VARIETY OF METABOLIC REACTIONS.  
 CC -1- COFACTOR: BINDS 1 2FE-2S CLUSTER.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL; M88251; AAA27329.1; -;  
 DR PIR; C47673; C47673.  
 DR HSSP; P27320; IDOX.  
 DR InterPro; IPR000564; 2FE2S.ferredoxln.  
 DR InterPro; IPR001041; Ferredoxin.  
 DR Pfam; PF00111; fer2. 1.  
 DR PRINTS; PR00159; 2FE2SERDOXIN.  
 DR PROSITE; PS00197; 2FE2S\_FERREDOXIN; 1.  
 KW Electron transport; Iron-sulfur.  
 FT INIT MET 0  
 FT METAL 39 39  
 FT METAL 44 44  
 FT METAL 47 47  
 FT METAL 77 77  
 FT SEQUENCE 96 AA; 10258 MW; 71F366CA9DF42549 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 96;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 DB 15 SYDA 18

## RESULT 4

PER\_LOXAL STANDARD; PRT; 109 AA.  
 ID PER\_LOXAL  
 AC Q25221;

DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Period circadian protein (Fragment).  
 GN PER.

OS Loxocera albispata (Rust fly).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 OC Muscomorpha; Diopsioidea; Psilidae; Psilinae; Loxocera.  
 NCBI\_Taxid=34685;

SEQUENCE FROM N.A.  
 RX MEDLINE=95115533; PubMed=7815924;  
 RA Nielsen J., Peixoto A.A., Piccin A., Costa R., Kyriacou C.P.,  
 RA Chalmers D.;  
 RT "Big files, small repeats: the 'Thr-Gly' region of the period gene in  
 RT Diptera.";

```

RL M01. Biol. Evol. 11:839-853(1994).
CC -1- FUNCTION: ESSENTIAL FOR BIOLOGICAL CLOCK FUNCTIONS. DETERMINES THE
CC PERIOD LENGTH OF CIRCADIAN AND ULTRADIAN RHYTHMS; AN INCREASE IN
CC PERIOD LENGTH LEADS TO SHORTENED CIRCADIAN RHYTHMS AND A DECREASE IN
CC LEADS TO LENGTHENED CIRCADIAN RHYTHMS. ESSENTIAL FOR THE CIRCADIAN
CC RHYTHMICITY OF LOCOMOTOR ACTIVITY, ECLOSION BEHAVIOR, AND FOR THE
CC RHYTHMIC COMPONENT OF THE MALE COURTSHIP SONG THAT ORIGINATES IN
CC THE THORACIC NERVOUS SYSTEM. THE BIOLOGICAL CYCLE DEPENDS ON THE
CC RHYTHMIC FORMATION AND NUCLEAR LOCALIZATION OF THE TIM-PER
CC COMPLEX. LIGHT INDUCES THE DEGRADATION OF TIM, WHICH PROMOTES
CC ELIMINATION OF PER. NUCLEAR ACTIVITY OF THE HETERODIMER
CC COORDINATIVELY REGULATES PER AND TIM TRANSCRIPTION THROUGH A
CC NEGATIVE FEEDBACK LOOP. BEHAVES AS A NEGATIVE ELEMENT IN CIRCADIAN
CC TRANSCRIPTIONAL LOOP. DOES NOT APPEAR TO BIND DNA, SUGGESTING
CC INDIRECT TRANSCRIPTIONAL INHIBITION (BY SIMILARITY).
CC -1- SUBUNIT: FORMS HETERODIMER WITH TIMELESS (TIM); THE COMPLEX THEN
CC TRANSLLOCATES INTO THE NUCLEUS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR AT SPECIFIC PERIODS OF THE DAY.
CC FIRST ACCUMULATES IN THE PERINUCLEAR REGION ABOUT ONE HOUR BEFORE
CC TRANSLLOCATION INTO THE NUCLEUS. INTERACTION WITH TIM IS REQUIRED
CC FOR NUCLEAR LOCALIZATION (BY SIMILARITY).
CC -1- PTM: PHOSPHORYLATED WITH A CIRCADIAN RHYTHMICITY, PROBABLY BY THE
CC DOUBLE-TIME PROTEIN (DBT). PHOSPHORYLATION COULD BE IMPLICATED IN
CC THE STABILITY OF PER MONOMER AND IN THE FORMATION OF HETERODIMER
CC PER-TIM (BY SIMILARITY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U11806; AAA76593.1; -
KW Biological rhythms; Repeat; Nuclear protein; Phosphorylation.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 109 AA; 11600 MW; D70555ABE9FC1699 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 109;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
   ||||
Db 43 SYDA 46

RESULT 5
FLGCG_AGR5 STANDARD; PRT; 139 AA.
ID FLGCG_AGR5
AC Q44336;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Flagellar basal-body rod protein FlgC.
DE FlgC OR ATU0554 OR AGR_C_975.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97305158; Pubmed=9161424;
RA Deakin W.J., Furniss C.S., Parker V.E., Shaw C.H.;
RT "Isolation and characterisation of a linked cluster of genes from
RT Agrobacterium tumefaciens encoding proteins involved in flagellar
RT basal-body structure."
RL Gene 189:135-137(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; Pubmed=11743193;

```

```

RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Kap P.D., Bovee D. Sr.,
RA Chapman P., Glendinning J., Deatherage G., Giller W., Grant C.,
RA Kutayav N.T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nestor E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RL Science 294:2317-2323(2001).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; Pubmed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Gourello B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmlet K., Gordon J., Vaudin M., Iarchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughy D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gurtson J., Lomo C., Sear C., Strud G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RL Science 294:2323-2328(2001).
CC -1- SUBUNIT: THE BASAL BODY CONSTITUTES A MAJOR PORTION OF THE
CC FLAGELLAR ORGANELLE AND CONSISTS OF FOUR RINGS (L,P,S, AND M)
CC MOUNTED ON A CENTRAL ROD. THE ROD CONSISTS OF ABOUT 26 SUBUNITS OF
CC FLAG IN THE DISTAL PORTION, AND FLGB, FLGC AND FLGF ARE THOUGHT TO
CC BUILD UP THE PROXIMAL PORTION OF THE ROD WITH ABOUT 6 SUBUNITS
CC EACH (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE FLAGELLA BASAL BODY ROD PROTEINS
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U39941; AAB68966.1; -
DR EMBL; U95165; AAB71788.1; -
DR EMBL; AE009024; AAL4571.1; -
DR EMBL; AE007990; AAK86366.1; -
DR InterPro: IPR001444; Flag_db-rod.
DR PROSITE: PS00588; FLAGELLA_BB-ROD; 1.
KW Flagella; Complete proteome.
FT CONFLICT 10 11
FT CONFLICT 101 101
FT CONFLICT 101 101
SQ SEQUENCE 139 AA; 15100 MW; F94340FA0C61D6B8 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 139;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
   ||||
Db 113 SYDA 116

RESULT 6
Y589_HAEIN STANDARD; PRT; 139 AA.
ID Y589_HAEIN
AC P44020;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein HI0589.
DE HI0589.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;

```

```

OC Haemophilus.
OX NCBI_Taxid=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghagen N.S.M.,
RA Gehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (potential).
CC -1- SIMILARITY: TO E.COLI RSEC AND H.INFLUENZAE H1850.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sib.ch).
CC -----
DR EMBL: U37740; AAC22246.1; -.
DR TIGR: H10589; -.
KW Hypothetical protein; Transmembrane; Inner membrane;
KW Complete proteome.
FT TRANSMEM 71 POTENTIAL.
FT TRANSMEM 97 117
SQ SEQUENCE 139 AA; 15635 MW; 6C26821FAF4DF32D CRC64;

Query Match 100.0%; Score 21; DB 1; Length 139;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
Db 3 SYDA 6

RESULT 7
COFI_YEAST
ID COFI_YEAST STANDARD; PRT; 143 AA.
AC 003048;
DR 01-JUL-1993 (Rel. 26, Created)
DR 01-JUL-1993 (Rel. 26, Last sequence update)
DR 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cofilin.
GN COFI OR YIL050C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_Taxid=4932;
RN [1]
RP SEQUENCE FROM N.A. AND SEQUENCE OF 43-56; 83-98 AND 106-141.
RX MEDLINE=93132073; PubMed=8421056;
RA Moon A.L., Jamney P.A., Louie K.A., Drubin D.G.;
RT "Cofilin is an essential component of the yeast cortical
RT cytoskeleton".
RL J Cell Biol. 120:421-435(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93178959; PubMed=8440472;
RA Iida K., Moriyama K., Matsumoto S., Kawasaki H., Nishida E.,
RA Yahara I.;
RT "Isolation of a yeast essential gene, COFI, that encodes a homologue

```

```

RT of mammalian cofilin, a low-M(r) actin-binding and depolymerizing
RT protein.";
RT Gene 124:115-120(1993).
RN [3]
RP SEQUENCE FROM N.A.
RA Wedler H., Wedler E., Scharfe M., Wambutt R.;
RL Submitted (May-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP INTERACTION WITH AIP1.
RX PubMed=10366597;
RA Rodal A.A., Tetreault J.W., Lappalainen P., Drubin D.G., Amberg D.C.;
RT "Aip1 interacts with cofilin to disassemble actin filaments.";
RL J. Cell Biol. 145:1251-1264(1999).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=97290449; PubMed=9145106;
RA Fedorov A.A., Lappalainen P., Fedorov E.V., Drubin D.G., Almo S.C.;
RT "Structure determination of yeast cofilin.";
RL Nat. Struct. Biol. 4:366-369(1997).
CC -1- FUNCTION: CONTROLS REVERSIBLY ACTIN POLYMERIZATION AND
CC DEPOLYMERIZATION IN A PH-SENSITIVE MANNER. IT HAS THE ABILITY TO
CC BIND G- AND F-ACTIN IN A 1:1 RATIO OF COFILIN TO ACTIN. IT IS THE
CC MAJOR COMPONENT OF INTRANUCLEAR AND CYTOPLASMIC ACTIN RODS. IN
CC EFFECT, YEAST COFILIN INCREASES THE RATE OF ACTIN POLYMERIZATION
CC BY MAKING NEW ENDS AVAILABLE FOR ACTIN SUBUNIT ADDITION. SUCH A
CC PROTEIN COMPLEX IS IMPORTANT FOR THE POLARIZED GROWTH OF YEAST
CC CELLS.
CC -1- SUBUNIT: Interacts with actin and AIP1 in a ternary complex.
CC -1- SUBCELLULAR LOCATION: THROUGHOUT THE CYTOPLASM (BUT NOT ON THE
CC CYTOPLASMIC CABLES) AND MAJOR COMPONENT OF THE CORTICAL ACTIN
CC CYTOSKELETON.
CC -1- PTM: THE N-TERMINUS IS BLOCKED.
CC -1- SIMILARITY: BELONGS TO THE ACTIN-BINDING PROTEINS ADF FAMILY.
CC SIMILAR TO THE N-TERMINUS OF YEAST ABPI PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sib.ch).
CC -----
DR EMBL: Z14971; CAAT8694.1; -.
DR EMBL: D13230; BAA02514.1; -.
DR EMBL: Z73155; CAA97502.1; -.
DR PIR: A44397; A44397.
DR PIR: S31309; S31309.
DR PIR: JN0529; JN0529.
DR PDB: 1COF; 01-APR-97.
DR PDB: 1CEY; 01-APR-97.
DR PDB: 1OPV; 08-JUN-99.
DR SGD: S0003973; COFI.
DR InterPro: IPR002108; Actbind_cofin.
DR Pfam: PF00241; cofilin_ADF. 1.
DR ProDom: PD002129; Actbind_cofin. 1.
DR SMART: SM00102; ADF. 1.
DR PROSITE: PS00325; ACTIN_DEPOLYMERIZING. 1.
KW Actin-binding; Cytoskeleton; 3D-structure.
FT DOMAIN 90 109 ACTIN-BINDING (POTENTIAL).
SQ SEQUENCE 143 AA; 15901 MW; 7A03747B0F21F22D CRC64;

Query Match 100.0%; Score 121; DB 1; Length 143;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
Db 49 SYDA 52

RESULT 8
SOD1_HAIME

```

```

ID SOD1_HALME STANDARD; PRT; 144 AA.
AC 008461;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Superoxide dismutase [Mn] 1 (EC 1.15.1.1) (Fragment).
GN SOD1.
OS Halobacterium mediterranei (Halobacter mediterranei).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Haloferax.
OX NCBI_TaxID=2252;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97257269; PubMed=9106366;
RA Dennis P.P., Shimmin L.C.;
RT "Evolutionary divergence and salinity-mediated selection in halophilic
archaea."
RL Microbiol. Mol. Biol. Rev. 61:90-104(1997).
CC -1- FUNCTION: Destroys toxic which are normally produced within the
cells and which are toxic to biological systems.
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) -> O(2) + H(2)O(2).
CC -1- COFACTOR: Manganese (By similarity).
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U78908; AAB60932.1; -
DR HSSP; P80857; ISSS.
DR InterPro; IPR001189; SODismutase.
DR Pfam; PF000081; sodfe; 1.
DR Pfam; PF02777; sodfe.C; 1.
DR ProDom; PD000475; SODismutase; 1.
DR PROSITE; PS00088; SOD_MN; 1.
KM Oxidoreductase; Manganese.
FT NON_TER 1
FT METAL 42 42 MANGANESE LIGAND (BY SIMILARITY).
FT METAL 124 124 MANGANESE LIGAND (BY SIMILARITY).
FT METAL 128 128 MANGANESE LIGAND (BY SIMILARITY).
FT NON_TER 144 144
SQ SEQUENCE 144 AA; 15613 MW; 47BBA0AC551A09A4 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 144;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
ID 1111
DB 71 SYDA 74

RESULT 9
GALX_CANMA STANDARD; PRT; 153 AA.
AC P56600;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE GAL10 bifunctional protein [Includes: UDP-glucose 4-epimerase
(EC 5.1.3.2) (Galactowaldenase); Aldose 1-epimerase (EC 5.1.3.3)
(Mutarotase)] (Fragment).
GN GAL10.
OS Candida maltosa (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5479;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX STRAIN=IAM 12247;
RX MEDLINE=97197967; PubMed=9046083;
RA Park S.M., Ohkuma M., Masuda Y., Ohta A., Takagi M.;
RT "Galactose-inducible expression systems in Candida maltosa using
RT promoters of newly-isolated GAL1 and GAL10 genes."
RL Yeast 13:21-29(1997).
CC -1- FUNCTION: Mutarotase converts alpha-aldose to the beta-anomer. It
is active on D-glucose, L-arabinose, D-xylose, D-galactose,
CC maltose and lactose (By similarity).
CC -1- CATALYTIC ACTIVITY: UDP-glucose = UDP-galactose.
CC -1- CATALYTIC ACTIVITY: Alpha-D-glucose = beta-D-glucose.
CC -1- COFACTOR: NAD.
CC -1- PATHWAY: Galactose metabolism; third step.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE SUGAR
EPIMERASE FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ALDOSE
EPIMERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D29759; -; NOT_ANNOTATED_CDS.
DR HSSP; P09147; 1XEL.
DR InterPro; IPR001823; Ald1 epimerase.
DR InterPro; IPR001509; Epimerase_Dh.
DR InterPro; IPR000205; NAD_binding.
DR InterPro; IPR001230; Prenyl_site.
DR Pfam; PF01370; Epimerase; 1.
DR PROSITE; PS00545; ALDOSE_1_EPIMERASE; PARTIAL.
KM Multifunctional enzyme; Isomerase; NAD; Galactose metabolism.
FT DOMAIN 1 >153 GALACTOMALDENASE.
FT NON_TER 153
SQ SEQUENCE 153 AA; 16752 MW; A069AC3040BC978D CRC64;

Query Match 100.0%; Score 21; DB 1; Length 153;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
ID 1111
DB 40 SYDA 43

RESULT 10
RIB2_PHOLE STANDARD; PRT; 156 AA.
AC Q93E92;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE 6,7-dimethyl-8-ribitylthiazine synthase (EC 2.5.1.9) (DMRL synthase)
DE (lumazine synthase) (Riboflavin synthase beta chain).
GN RIBH.
OS Photobacterium leiognathi.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae;
OC Photobacterium.
OX NCBI_TaxID=658;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=PL741;
RX MEDLINE=21290647; PubMed=11396941;
RA Lin J.-W., Chao Y.-F., Weng S.-F.;
RT "Riboflavin synthesis genes ribe, ribp, ribb, riba reside in the lux
RT operon of Photobacterium leiognathi."
RL Biochem. Biophys. Res. Commun. 284:587-595(2001).
CC -1- FUNCTION: Riboflavin synthase is a bifunctional enzyme complex

```

```

CC catalyzing the formation of riboflavin from 5-amino-6-(1'-D)-
CC ribityl-amino-2,4(1H,3H)-pyrimidinedione and L-3,4-dihydroxy-2-
CC butanone-4-phosphate via 6,7-dimethyl-8-lumazine. The beta subunit
CC catalyzes the condensation of 5-amino-6-(1'-D)-ribityl-amino-
CC 2,4(1H,3H)-pyrimidinedione with L-3,4-dihydroxy-2-butanone-4-
CC phosphate yielding 6,7-dimethyl-8-lumazine (By similarity).
CC CATALYTIC ACTIVITY: 2 6,7-dimethyl-8-(1-D-ribityl)lumazine =
CC riboflavin + 4-(1-D-ribitylamino)-5-amino-2,6-dihydroxypyrimidine.
CC -1- SIMILARITY: BELONGS TO THE DMRL SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF364106; AAK83294.1; -
DR InterPro: IPR002180; DMRL_synthase.
DR Pfam: PF00885; DMRL_synthase; 1.
DR ProDom: PD003664; DMRL_synthase; 1.
DR TIGRFAMs: TIGR00114; ribH.1.
KM Riboflavin biosynthesis; transferase.
SQ SEQUENCE 156 AA; 16423 MW; B53E2727C6BCD0A5 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 156;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
    ||||
Db 72 SYDA 75

RESULT 11
RL2B_TRYBB
ID RL2B_TRYBB STANDARD; PRT; 164 AA.
AC P41165;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DE 60S ribosomal protein L23a (L25).
GN RPL23A OR RPL25.
OS Trypanosoma brucei brucei.
OC Eukaryota; Euzoenzoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5702;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-ISTAT 1.1;
RX MEDLINE=94232821; PubMed=8177742;
RA Metzenberg S., Jodel C., Verspielen P., Agabian N.;
RT "Ribosomal protein L25 from Trypanosoma brucei: phylogeny and
RT molecular co-evolution of an rRNA-binding protein and its rRNA
RT binding site.";
RL Nucleic Acids Res. 21:4936-4940(1993).
-1- FUNCTION: THIS PROTEIN BINDS TO A SPECIFIC REGION ON THE 26S RNA
-1- (BY SIMILARITY).
-1- SIMILARITY: BELONGS TO THE L23P FAMILY OF RIBOSOMAL PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L21172; AAC37186.1; -
DR InterPro: IPR001014; Ribosomal_L23.
DR Pfam: PF00276; Ribosomal_L23; 1.
DR ProDom: PD001141; Ribosomal_L23; 1.

```

```

DR PROSITE: PS00050; RIBOSOMAL_L23; 1.
KW Ribosomal protein; rRNA-binding.
SQ SEQUENCE 164 AA; 18170 MW; 1606F4E8BFC0C8F CRC64;

Query Match 100.0%; Score 21; DB 1; Length 164;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
    ||||
Db 151 SYDA 154

RESULT 12
HXD3_BRARE
ID HXD3_BRARE STANDARD; PRT; 167 AA.
AC 042370;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-D3 (Fragment).
GN HOXD3 OR HOXD3A.
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN (1)
RP SEQUENCE FROM N.A.
RA Prince V.E., Moens C.B., Kimmel C.B., Ho R.K.;
RT "Zebrafish hox genes: expression in the hindbrain region of wild-type
RT and mutants of the segmentation gene, valentino.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
-1- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF
CC A DEVELOPMENTAL REGULATORY SYSTEM THAT PROVIDES CELLS WITH
CC SPECIFIC POSITIONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE ANTP HOMEOBOX FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y13948; CAA74286.1; -
DR TRANSFAC: T03651; -
DR ZFIN: ZDB-GENE-990415-120; hoxd3a.
DR InterPro: IPR000047; HTH_repressr.
DR ProDom: PD000047; HTH_repressr.
DR Pfam: PF00046; homeobox; 1.
DR PRINTS: PR00024; HOMEOBOX.
DR PROSITE: PS00031; HTHREPRESSR.
DR ProDom: PD000010; Homeobox; 1.
DR SMART: SM00389; HOX.1.
DR PROSITE: PS00027; HOMEOBOX_1; 1.
DR PROSITE: PS00071; HOMEOBOX_2; 1.
KW Homeobox; DNA-binding; Developmental
KW Transcription regulation.
FT NON_TER 1
FT DNA_BIND <1 32 HOMEOBOX.
FT NON_TER 167
SQ SEQUENCE 167 AA; 18339 MW; E90327377FF865CF CRC64;

Query Match 100.0%; Score 21; DB 1; Length 167;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
    ||||

```

Db 71 SYDA 74

RESULT 13

SCYD\_ASPFU STANDARD; PRT; 168 AA.

AC 014434;

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Probable scytalone dehydratase (EC 4.2.1.94).

GN ARPI.

OS *Aspergillus fumigatus* (Sartorya fumigata).

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; *Aspergillus*.

OX NCBI\_TaxID=5085;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=B-5233;

RX MEDLINE=98043547; PubMed=9383199;

RA Tsai H.-F., Washburn R.G., Chang Y.C., Kwon-Chung K.J.;

RT *Aspergillus fumigatus* arpi modulates conidial pigmentation and complement deposition.

RL Mol. Microbiol. 26:175-183(1997).

CC -1- FUNCTION: MODULATES THE BLuish-GREEN PIGMENTATION OF CONIDIA AS WELL AS COMPLEMENT DEPOSITION.

CC -1- CATALYTIC ACTIVITY: scytalone = 1,3,8-trihydroxynaphthalene + H(2)O.

CC -1- PATHWAY: BIOSYNTHESIS OF FUNGAL BLuish-GREEN PIGMENT.

CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING CONIDIATION.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: U95042; AAC49843.1; -.

DR HSSP: P56221; 3STD.

DR InterPro: IPR004235; Scytalone\_DH.

DR Pfam: PF02982; Scytalone\_DH; 1.

KW Lyase; Conidiation.

CC SEQUENCE 168 AA; 19716 MW; F40C22CEB6D02695 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 168;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4

Db 28 SYDA 31

RESULT 14

RBS2\_MESCR STANDARD; PRT; 180 AA.

AC 004450;

DT 01-OCT-1994 (Rel. 30, Created)

DT 01-OCT-1994 (Rel. 30, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Ribulose biphosphate carboxylase small chain 2, chloroplast precursor (EC 4.1.1.39) (Rubisco small subunit 2).

GN RBCS-2.

OS *Mesembryanthemum crystallinum* (Common ice plant).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Caryophyllidae; Caryophyllales; Alzooceae; *Mesembryanthemum*.

OX NCBI\_TaxID=3544;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=93302719; PubMed=8316216;

RA Derocher E.J., Quigley F., Mache R., Bohnert H.J.;

RT "The six genes of the Rubisco small subunit multigene family from *Mesembryanthemum crystallinum*, a facultative CAM plant.;

RL Mol. Gen. Genet. 239:450-462(1993).

RN [2]

RP SEQUENCE FROM N.A.

RA Derocher E.J., Michalowski C.B., Bohnert H.J.;

RT "cDNA sequences for transcripts of the ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit gene family of *Mesembryanthemum crystallinum*.";

RL Plant Physiol. 95:976-978(1991).

CC -1- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF D-RIBULOSE 1,5-BISPHOSPHATE. THE PRIMARY EVENT IN PHOTOSYNTHETIC CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF THE PENTOSE SUBSTRATE IN THE PHOTOSPIRATION PROCESS. BOTH REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME ACTIVE SITE.

CC -1- CATALYTIC ACTIVITY: D-ribulose 1,5-bisphosphate + CO(2) = 2 3-phospho-D-glycerate.

CC -1- CATALYTIC ACTIVITY: D-ribulose 1,5-bisphosphate + O(2) = 3-phospho-D-glycerate + 2-phosphoglycolate.

CC -1- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.

CC -1- SIMILARITY: BELONGS TO THE RUBISCO SMALL CHAIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: L10214; AAA03694.1; -.

DR EMBL: M88316; AAA33036.1; -.

DR PIR: S35246; S35246.

DR InterPro: IPR000894; Rubisco\_small.

DR Pfam: PF00101; Rubisco\_small; 1.

DR PRINTS: PR00152; RUBISCO\_SMALL.

DR ProDom: PD000290; Rubisco\_small; 1.

KW Photosynthesis; Carbon dioxide fixation; Photorespiration;

KW Lyase; Oxidoreductase; Monooxygenase; Chloroplast; Transil peptide;

KW Multigene family.

FT TRANSIT 1 55 CHLOROPLAST (BY SIMILARITY).

FT CHAIN 56 180 RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL

FT CHAIN 2.

FT CONFLICT 154 154 T -> I (IN REF. 2).

SO SEQUENCE 180 AA; 20084 MW; 6977DC72C89940F3 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 180;

Best Local Similarity 100.0%; Pred. No. 1.5e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4

Db 177 SYDA 180

RESULT 15

RBS1\_MESCR STANDARD; PRT; 182 AA.

AC P16032;

DT 01-APR-1990 (Rel. 14, Created)

DT 01-OCT-1994 (Rel. 30, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Ribulose biphosphate carboxylase small chain 1, chloroplast precursor (EC 4.1.1.39) (Rubisco small subunit 1).

GN RBCS-1 OR RBC5 OR RBCS.

OS *Mesembryanthemum crystallinum* (Common ice plant).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Caryophyllidae; Caryophyllales; Alzooceae; *Mesembryanthemum*.

OX NCBI\_TaxID=3544;

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93302719; PubMed=8316216;
RA Derocher E.J., Outigley F., Mache R., Bohnert H.J.;
RT "The six genes of the Rubisco small subunit multigene family from
RL Mesembryanthemum crystallinum, a facultative CAM plant.";
RN Mol. Gen. Genet. 239:450-462(1993).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87316892; PubMed=3627990;
RA Derocher E.J., Ramage R.T., Michalowski C.B., Bohnert H.J.;
RT "Nucleotide sequence of a cDNA encoding rbcS from the desert plant
RL Mesembryanthemum crystallinum.";
RN Nucleic Acids Res. 15:6301-6301(1987).
CC -1- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF
CC D-RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC
CC CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
CC THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
CC REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
CC ACTIVE SITE.
CC -1- CATALYTIC ACTIVITY: D-ribulose 1,5-bisphosphate + CO(2) = 2 3-
CC phospho-D-glycerate.
CC -1- CATALYTIC ACTIVITY: D-ribulose 1,5-bisphosphate + O(2) =
CC 3-phospho-D-glycerate + 2-phosphoglycolate.
CC -1- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- SIMILARITY: BELONGS TO THE RUBISCO SMALL CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L10212; AAA03693.1; -
DR EMBL; M31640; AAA33035.1; -
DR PIR; A29118; RKXS
DR PIR; S35247; S35247.
DR InterPro; IPR000694; RUBISCO_small.
DR Pfam; PF00101; RUBISCO_small.1.
DR PRINTS; PR00152; RUBISCO_SMALL.
DR ProDom; PD000290; RUBISCO_small.1.
DR KMW; Photosynthesis: Carbon dioxide fixation; Photorespiration;
DR KMW; Lyase; Oxidoreductase; Monooxygenase; Chloroplast; Transit peptide;
DR KMW; Multigene family.
FT TRANSIT 1 57 CHLOROPLAST (BY SIMILARITY).
FT CHAIN 1 58 182 RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL
FT CHAIN 1.
FT CONFLICT 25 25 A -> V (IN REF. 2).
FT CONFLICT 145 145 E -> R (IN REF. 2).
SQ SEQUENCE 182 AA; 20276 MW; 7AB023654AF05665 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 182;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
| | | |
DB 179 SYDA 182

RESULT 16
ATPD_BACCHD STANDARD; PRT; 183 AA.
AC 09K6H2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE ATP synthase delta chain (EC 3.6.3.14).
GN ATPH OR BH3757.
OS Bacillus halodurans.

```

```

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxId=86665;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y.,
RA Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RL halodurans and genomic sequence comparison with Bacillus subtilis.";
RN Nucleic Acids Res. 28:4317-4331(2000).
CC -1- FUNCTION: THIS PROTEIN SEEMS TO BE PART OF THE STALK THAT LINKS
CC CF(0) TO CF(1). IT EITHER TRANSMITS CONFORMATIONAL CHANGES FROM
CC CF(0) INTO CF(1) OR IS IMPLICATED IN PROTON CONDUCTION (BY
CC SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +
CC H(+) (out).
CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
CC HAS THREE MAIN SUBUNITS: A, B AND C.
CC -1- SIMILARITY: BELONGS TO THE ATPASE DELTA CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP001519; BAB07476.1; -
DR InterPro; IPR000711; ATPsynth_OSCP.
DR Pfam; PF00213; OSCP.1.
DR PRINTS; PR00125; ATPASEDELTA.
DR TIGRFAMS; TIGR01145; ATP_synth_delta.1.
DR PROSITE; PS00389; ATPASE_DELTA.
DR KMW; Hydrolyase; ATP synthetase; CF(1); Hydrogen ion transport;
DR KMW; Complete proteome.
SQ SEQUENCE 183 AA; 20605 MW; 5CB7989AFA5D26F2 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 183;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
| | | |
DB 101 SYDA 104

RESULT 17
RBS3_MESCR STANDARD; PRT; 183 AA.
AC 008183;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ribulose biphosphate carboxylase small chain 3, chloroplast precursor
DE (EC 4.1.1.39) (Rubisco small subunit 3).
GN RBCS-3.
OS Mesembryanthemum crystallinum (Common ice plant).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Aizoaceae; Mesembryanthemum.
OX NCBI_TaxId=3544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93302719; PubMed=8316216;
RA Derocher E.J., Outigley F., Mache R., Bohnert H.J.;
RT "The six genes of the Rubisco small subunit multigene family from
RL Mesembryanthemum crystallinum, a facultative CAM plant.";
RN Mol. Gen. Genet. 239:450-462(1993).

```



[2]  
 RN SEQUENCE FROM N.A.  
 RA Derocher E.J., Michalowski C.B., Bohnert H.J.;  
 RT "cDNA sequences for transcripts of the ribulose-1,5-bisphosphate  
 carboxylase/oxygenase small subunit gene family of Mesembryanthemum  
 crystallinum";  
 RL Plant Physiol. 95:976-978 (1991).  
 CC -1- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF  
 D-RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC  
 CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF  
 THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH  
 REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME  
 ACTIVE SITE.  
 CC -1- CATALYTIC ACTIVITY: D-ribulose 1,5-bisphosphate + CO(2) = 2 3-  
 phospho-D-glycerate.  
 CC -1- CATALYTIC ACTIVITY: D-ribulose 1,5-bisphosphate + O(2) =  
 3-phospho-D-glycerate + 2-phosphoglycolate.  
 CC -1- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.  
 CC -1- SIMILARITY: BELONGS TO THE RUBISCO SMALL CHAIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: L10215; AAA03695.1; -;  
 DR EMBL: M38317; AAA33037.1; -;  
 DR PIR: S35245; S35245.  
 DR InterPro: IPR000894; Rubisco\_small.  
 DR Pfam: PF00101; Rubisco\_small; 1.  
 DR PRINTS: PR00152; RUBISCO\_SMALL.  
 DR ProDom: PD000290; Rubisco\_small; 1.  
 DR KMW Photosynthesis; Carbon dioxide fixation; Photorespiration;  
 KMW Lyase; Oxidoreductase; Monooxygenase; Chloroplast; Transit peptide;  
 KMW Multigene family.  
 FT TRANSIT 1 58 CHLOROPLAST (BY SIMILARITY).  
 FT CHAIN 59 183 RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL  
 FT CONFIDENT 85 85 CHAIN 3.  
 FT SEQUENCE 183 AA; 20385 MW; F056MAA5B110AA71 CRC64;  
 SQ  
 Query Match 100.0%; Score 21; DB 1; Length 183;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 SYDA 4  
 DB 180 SYDA 183  
 RESULT 18  
 RBS6\_MESCR STANDARD; PRT; 186 AA.  
 AC 008186;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Ribulose biphosphate carboxylase small chain 6, chloroplast precursor  
 DE (EC 4.1.1.39) (Rubisco small subunit 6).  
 GN RBCS-6.  
 OS Mesembryanthemum crystallinum (Common ice plant).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Caryophyllidae; Caryophyllales; Alzaceae; Mesembryanthemum.  
 OX NCBI\_TaxID=3544;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93302719; PubMed=8316216;  
 RA Derocher E.J., Quigley F., Mache R., Bohnert H.J.;

RT "The six genes of the Rubisco small subunit multigene family from  
 Mesembryanthemum crystallinum, a facultative CAM plant.";  
 RL Mol. Gen. Genet. 239:450-462(1993).  
 CC -1- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF  
 D-RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC  
 CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF  
 THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH  
 REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME  
 ACTIVE SITE.  
 CC -1- CATALYTIC ACTIVITY: D-ribulose 1,5-bisphosphate + CO(2) = 2 3-  
 phospho-D-glycerate.  
 CC -1- CATALYTIC ACTIVITY: D-ribulose 1,5-bisphosphate + O(2) =  
 3-phospho-D-glycerate + 2-phosphoglycolate.  
 CC -1- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.  
 CC -1- SIMILARITY: BELONGS TO THE RUBISCO SMALL CHAIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: L10213; AAA03698.1; -;  
 DR PIR: S35242; S35242.  
 DR InterPro: IPR000894; Rubisco\_small.  
 DR Pfam: PF00101; Rubisco\_small; 1.  
 DR PRINTS: PR00152; RUBISCO\_SMALL.  
 DR ProDom: PD000290; Rubisco\_small; 1.  
 DR KMW Photosynthesis; Carbon dioxide fixation; Photorespiration;  
 KMW Lyase; Oxidoreductase; Monooxygenase; Chloroplast; Transit peptide;  
 KMW Multigene family.  
 FT TRANSIT 1 61 CHLOROPLAST (BY SIMILARITY).  
 FT CHAIN 62 186 RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL  
 FT SEQUENCE 186 AA; 20494 MW; 5825DF86F86B7F8 CRC64;  
 SQ  
 Query Match 100.0%; Score 21; DB 1; Length 186;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 SYDA 4  
 DB 183 SYDA 186  
 RESULT 19  
 SODF\_NICPL STANDARD; PRT; 202 AA.  
 AC P22302;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Superoxide dismutase [Fe], chloroplast (EC 1.15.1.1) (Fragment).  
 GN SODB.  
 OS Nicotiana glauca (Leadwort-leaved tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_TaxID=4092;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. P2;  
 RX MEDLINE=91086619; PubMed=2263641;  
 RA Van Camp W., Bowler C., Villarroel R., Tsang E.W.T., van Montagu M.,  
 Inze D.;  
 RT "Characterization of iron superoxide dismutase cDNAs from plants  
 obtained by genetic complementation in Escherichia coli.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:9903-9907(1990).  
 CC -1- FUNCTION: Destroys radicals which are normally produced within the  
 cells and which are toxic to biological systems.

```

CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -1- COFACTOR: Iron.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- INDUCTION: BY ENVIRONMENTAL CONDITIONS.
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: M55909; AAA34074.1; -.
CC PIR: A39267; A39267.
CC HSSP: P09157; IISC.
CC InterPro: IPR01189; SODismutase.
CC Pfam: PF00081; sodfe_1.
CC ProDom: PD000475; SODismutase; 1.
CC PROSITE: PS00088; SOD_MN; 1.
CC Oxidoreductase; Iron; Chloroplast.
CC NON_TER 1 1
CC METAL 26 26 IRON (BY SIMILARITY).
CC METAL 78 78 IRON (BY SIMILARITY).
CC METAL 162 162 IRON (BY SIMILARITY).
CC METAL 166 166 IRON (BY SIMILARITY).
CC SEQUENCE 202 AA; 23042 MW; 8659E79BDD0302C27 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 202;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4
    |||
Db 107 SYDA 110

RESULT 20
FLUID_PROMI STANDARD; PRT; 205 AA.
AC P42274;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Flagellar hook-associated protein 2 (HAP2) (Filament cap protein)
DE (Flagellar cap protein) (Fragment).
GN FLUID.
OS Proteus mirabilis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC NCBI_TaxID=564;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=BB2000;
RX MEDLINE=95011656; PubMed=7926835;
RA "Belas R., Flaherty D.;
RT "Sequence and genetic analysis of multiple flagellin-encoding genes
RT from Proteus mirabilis.";
RL Gene 148:33-41(1994).
-1- FUNCTION: REQUIRED FOR THE MORPHOGENESIS AND FOR THE ELONGATION OF
-1- THE FLAGELLAR FILAMENT BY FACILITATING POLYMERIZATION OF THE
-1- FLAGELLIN MONOMERS AT THE TIP OF GROWING FILAMENT. FORMS A CAPPING
-1- STRUCTURE, WHICH PREVENTS FLAGELLIN SUBUNITS (TRANSPORTED THROUGH
-1- THE CENTRAL CHANNEL OF THE FLAGELLUM) FROM LEAKING OUT WITHOUT
-1- POLYMERIZATION AT THE DISTAL END (BY SIMILARITY).
-1- SUBUNIT: HOMOPENTAMER (BY SIMILARITY).
-1- SUBCELLULAR LOCATION: Flagellar.
-1- SIMILARITY: BELONGS TO THE FLID FAMILY.
-----

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: AF221596; AAA62395.1; -.
CC InterPro: IPR003481; FLID.
CC Pfam: PF02465; FLID; 1.
CC Flagella.
CC NON_TER 205 205
CC SEQUENCE 205 AA; 22281 MW; 7BF8510388451416 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 205;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4
    |||
Db 39 SYDA 42

RESULT 21
FTSQ_STRGR STANDARD; PRT; 208 AA.
ID FTSQ_STRGR
AC P45503;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cell division protein ftsq homolog (Fragment).
GN FTSQ.
OS Streptomyces griseus.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1911;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B2682;
RX MEDLINE=94374704; PubMed=8088545;
RA Dharmatilake A.J., Kendrick K.E.;
RT "Expression of the division-controlling gene ftsz during growth and
RT sporulation of the filamentous bacterium Streptomyces griseus.";
RL Gene 147:21-28(1994).
-1- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN SEPTUM FORMATION (BY
-1- SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein (By similarity).
CC -1- SIMILARITY: BELONGS TO THE FTSQ FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: U07344; AAA56888.1; -.
CC Cell division; Septation; Transmembrane.
CC NON_TER 1 1
CC DOMAIN <1 208 EXTRACELLULAR (POTENTIAL).
FT SEQUENCE 208 AA; 21942 MW; 7CC9A86699746E82 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 208;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4
    |||
Db 153 SYDA 156

```

RESULT 22	HBX3_BRARE	STANDARD;	PRT;	210 AA.
AC	042368;			
DT	15-JUL-1999 (Rel. 38, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Homeobox protein Hox-B3 (Fragment).			
GN	HOXB3 OR HOXB3A.			
OS	Brachydanio rerio (Zebrafish) (Danio rerio).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;			
OC	Cyprinidae; Danio.			
OX	NCBI_Taxid=7955;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND DEVELOPMENTAL STAGE.			
RC	TISSUE=Embryo;			
RX	MEDLINE=98165394; PubMed=9425135;			
RA	Prince V.E., Moens C.B., Kimmel C.B., Ho R.K.;			
RT	"Zebrafish hox genes: expression in the hindbrain region of wild-type			
RT	and mutants of the segmentation gene, 'valentino.'";			
RL	Development 125:393-406(1998).			
RN	[2]			
RP	REVISTONS TO 79 AND C-TERMINUS.			
RC	TISSUE=Embryo;			
RA	Prince V.E.;			
RL	Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.			
CC	-1- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF			
CC	A DEVELOPMENTAL REGULATOR SYSTEM THAT PROVIDES CELLS WITH			
CC	SPECIFIC POSITIONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS (BY			
CC	SMILARITY).			
CC	-1- SUBCELLULAR LOCATION: Nuclear.			
CC	-1- DEVELOPMENTAL STAGE: EXPRESSED AT HIGH LEVELS IN RHOMBOMERES 5 AND			
CC	6 IN THE DEVELOPING HINDBRAIN DURING EMBRYOGENESIS.			
CC	-1- SIMILARITY: BELONGS TO THE ANTP HOMEOBOX FAMILY			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; Y13944; CAA74282.2; -.			
DR	TRANSFAC; T03631; -.			
DR	ZFIN; ZDB-GENE-990415-104; hoxb3a.			
DR	InterPro: IPR000047; HTH_repressr.			
DR	InterPro: IPR001356; Homeobox.			
DR	Pfam; PF00046; homeobox.1.			
DR	PRINTS; PR00024; HOMEOBOX.			
DR	PRINTS; PR00031; HTHREPRESSR.			
DR	ProDom; PD000010; Homeobox; 1.			
DR	SMART; SM00389; Hox; 1.			
DR	PROSITE; PS00027; HOMEOBOX_1; 1.			
DR	PROSITE; PS00077; HOMEOBOX_2; 1.			
KW	Homeobox; DNA-binding; Developmental protein; Nuclear protein;			
KW	Transcription regulation.			
FT	NON_TER	1	1	
FT	DNA_BIND	<1	32	HOMEOBOX.
FT	SEQUENCE	210 AA;	22862 MW;	55B0D2D2D79E8601 CRC64;

Query Match	100.0%	Score 21;	DB 1;	Length 210;
Best Local Similarity	100.0%;	Pred. No. 1.7e+02;		
Matches	4;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0

```

OY      1 SYDA 4
        ||||
Db      69 SYDA 72

```

ID	ATKC_HAINT	STANDARD	PRT	216 AA.
AC	P57687			
DT	16-OCT-2001	(Rel. 40, Created)		
DT	16-OCT-2001	(Rel. 40, Last sequence update)		
DT	16-OCT-2001	(Rel. 40, Last annotation update)		
DE	Potassium-transporting ATPase C chain (EC 3.6.3.12) [Potassium-transporting ATPase C chain] (ATP phosphotransferase [potassium-transporting] C chain) (Potassium binding and translocating subunit C).			
GN	KDPC OR VNG6178G.			
OS	Halobacterium sp. (strain NRC-1).			
OG	Plasmid pNRC200.			
OC	Archaea; Euryarchaeota; Halobacteriia; Halobacteriales;			
CC	Halobacteriaceae; Halobacterium.			
CC	NCBI_TaxID=64091.			
CC	SEQUENCE FROM N.A.			
CC	MEDLINE=20504483; PubMed=11016950;			
CC	Ng W.Y., Kennedy S.P., Mahairs G.G., Bergquist B., Pan M., Shukla H.D., Laaky S.R., Baliga N.S., Thorsson V., Sirogaa J., Swartzell S., Weir D., Hall J., Dahl T.A., Weltl R., Goox Y.A., Leithauer B., Keller K., Cruz R., Danson M.J., Hough D.W., Maddocks D.G., Jablonksi P.E., Krebs M.P., Angewine C.M., Dale H., Idenberger T.A., Peck R.F., Pohlschoder M., Spudich J.L., Jung K.-H., Alam M., Freltas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D., Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S., "Genome sequence of Halobacterium species NRC-1"; Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).			
CC	-1- FUNCTION: ONE OF THE COMPONENTS OF THE HIGH-AFFINITY ATP-DRIVEN POTASSIUM TRANSPORT (OR KDP) SYSTEM, WHICH CATALYZES THE HYDROLYSIS OF ATP COUPLED WITH THE EXCHANGE OF HYDROGEN AND POTASSIUM IONS. THE C SUBUNIT MAY BE INVOLVED IN ASSEMBLY OF THE KDP COMPLEX (BY SIMILARITY).			
CC	-1- CATALYTIC ACTIVITY: ATP + H(2)O + K(+) (Out) = ADP + phosphate + K(+) (In).			
CC	-1- SUBCELLULAR LOCATION: Membrane-associated (Probable).			
CC	-1- SIMILARITY: BELONGS TO THE KDPC FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	EMBL: AE005153; AAC30845.1; -			
DR	InterPro: IPR003820; K_ATPase_KDPC.			
DR	Pfam: PF02669; kdpc; 1.			
KW	Hydrolase; Transport; Potassium transport; Transmembrane; Plasmid; Complete proteome.			
FT	TRANSMEM 12 32 POTENTIAL.			
SO	SEQUENCE 216 AA; 23768 MW; 97616E5C59E6107 CRC64;			
QY	1 SYDA 4			
QY	1111			
Db	132 SYDA 135			
RESULT 24				
ID	OLII_HUMAN	STANDARD;	PRT;	216 AA.
AC	P47889			
DT	01-FEB-1996	(Rel. 33, Created)		
DT	01-FEB-1996	(Rel. 33, Last sequence update)		
DT	16-OCT-2001	(Rel. 40, Last annotation update)		
DE	Olfactory receptor-like protein OR11-207 (Fragment).			
OS	Homo sapiens (Human).			
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			

RESULT	24
OLII_HUMAN	
ID	OLII_HUMAN
AC	P47889;
DT	01-FEB-1996 (Rel. 33, Created)
DT	01-FEB-1996 (Rel. 33, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	Olfactory receptor-like protein OR17-207 (Fragment).
OS	Homo sapiens (Human).
OC	Eumaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=94272458; PubMed=8004088;
RA Ben-Arie N., Lancel D., Taylor C., Khen M., Walker N.,
RA Lebetter D.H., Carozzo R., Patel K., Sheer D., Lehnach H.,
RA North M.A.;
RT "Olfactory receptor gene cluster on human chromosome 17: possible
RT duplication of an ancestral receptor repertoire.";
RL Hum. Mol. Genet. 3:229-235(1994)
CC -1- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U04687; AAA18350.1; -
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC Pfam; PF00001; 7tm_1; 1.
CC PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
CC PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Multigene family; Olfaction.
FT NON_TER 1
FT TRANSMEM <1 12 2 (POTENTIAL).
FT DOMAIN 13 33 3 (POTENTIAL).
FT TRANSMEM 34 53 3 (POTENTIAL).
FT DOMAIN 54 72 2 (POTENTIAL).
FT TRANSMEM 74 91 4 (POTENTIAL).
FT TRANSMEM 92 129 5 (POTENTIAL).
FT TRANSMEM 130 173 5 (POTENTIAL).
FT DOMAIN 154 170 1 (POTENTIAL).
FT TRANSMEM 171 194 6 (POTENTIAL).
FT DOMAIN 195 205 7 (POTENTIAL).
FT TRANSMEM 206 >216 7 (POTENTIAL).
FT NON_TER 216
SQ SEQUENCE 216 AA; 23617 MW; 280BCBE39733F9F2 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 216;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4
Db 26 SYDA 29

RESULT 25
U119_CAEEL STANDARD: PRT; 217 AA.
AC Q17297;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Unc-119 protein.
GN UNC-119.
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_taxid=6238;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97149282; PubMed=8996090;
RA Maduro M.F., Pilgrim D.B.;
RT "Conservation of function and expression of unc-119 from two
RT Caenorhabditis species despite divergence of non-coding DNA.";
```

```

RL Gene 183:77-85(1996).
CC -1- FUNCTION: REQUIRED FOR THE ESTABLISHMENT OR FUNCTION OF THE
CC NERVOUS SYSTEM (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PDE6D / UNC119 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U45326; AAB41283.1; -
CC Developmental protein.
KW SEQUENCE 217 AA; 25094 MW; EACF90F99568BA67 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4
Db 214 SYDA 217

RESULT 26
U119_CAEEL STANDARD: PRT; 219 AA.
AC Q10658;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Unc-119 protein.
GN UNC-119 OR M142.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RX MEDLINE=96129273; PubMed=8582641;
RA Maduro M.F., Pilgrim D.B.;
RT "Identification and cloning of unc-119, a gene expressed in the
RT Caenorhabditis elegans nervous system.";
RL Genetics 141:977-988(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA McMurray A.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DDJ databases.
CC -1- FUNCTION: REQUIRED FOR THE ESTABLISHMENT OR FUNCTION OF THE
CC NERVOUS SYSTEM.
CC -1- TISSUE SPECIFICITY: PREDOMINANTLY NEURON-SPECIFIC.
CC -1- MISCELLANEOUS: MUTATIONS IN UNC-119 CAUSE DIVERSE NERVOUS SYSTEM
CC DEFECTS.
CC -1- SIMILARITY: BELONGS TO THE PDE6D / UNC119 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U32854; AAC46919.1; -
CC EMBL; 273428; CAA97807.1; -
CC Wormpep; M142.1; CE06203.
KW Developmental protein.
SQ SEQUENCE 219 AA; 25263 MW; 26F434575CDBB45 CRC64;
```

Query Match 100.0%; Score 21; DB 1; Length 219;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 Db 216 SYDA 219

## RESULT 27

Y945\_METJA STANDARD; PRT; 224 AA.  
 AC 058355;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein M0945.  
 GN M0945.  
 OS Methanococcus jannaschii.  
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
 OC Methanocaldococcaceae; Methanocaldococcus.  
 OX NCBI\_TaxID=2190;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JUL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE=96337999; PubMed=6680887;  
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
 RA Kerlavage A.R., Dougherty B.A., Tomb M.D., Reisch C.I.,  
 RA Overbeek R., Kirkness E.F., Weisslock K.G., Merrick J.M., Glodek A.,  
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,  
 RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
 RA Cotton M.D., Roberts K.M., Hurt M.A., Kaine B.P., Borodovsky M.,  
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
 RA Complete genome sequence of the methanogenic archaeon, Methanococcus  
 RT jannaschii.";  
 RL Science 273:1058-1073(1996).

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL: U67538; AAB98955.1; -;  
 DR TIGR: M0945;  
 KM Hypothetical protein: Transmembrane; Complete proteome.  
 FT TRANSMEM 21 41 POTENTIAL.  
 SQ SEQUENCE 224 AA; 25416 MW; 1E9F7E32C025DA9 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 224;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 Db 160 SYDA 163

## RESULT 28

Y9VA\_METTF STANDARD; PRT; 227 AA.  
 AC P29377;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 01-APR-1993 (Rel. 25, Last annotation update)  
 DE Hypothetical 24.7 kDa protein (ORF5A).  
 OS Methanobacterium thermoformicicum.  
 OC Plasmid PFV1.  
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;  
 OC Methanobacteriaceae; Methanothermobacter.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

OX NCBI\_TaxID=145262;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-DSM 3848 / TFE;  
 RA MEDLINE=93126090; PubMed=1336177;  
 RA Noelling J., van Eeden F.J.M., Eggen R.I.L., de Vos W.M.;  
 RT "Modular organization of related Archaeal plasmids encoding different  
 RT restriction/modification systems in Methanobacterium  
 RT thermoformicicum.";  
 RT Nucleic Acids Res. 20:6501-6507(1992).  
 RL Nucleic Acids Res. 20:6501-6507(1992).  
 CC -I- SIMILARITY: TO ORF5 IN PFZ1.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: X68366; CAA48429.1; -;  
 DR PIR: S26440; S26440.  
 DR PIR: S30305; S30305.  
 DR InterPro: IPR001434; DUF11.  
 DR Pfam: PF01345; DUF11.1.  
 KW Hypothetical protein. Plasmid.  
 SQ SEQUENCE 227 AA; 24681 MW; 2F3004C2519E849E CRC64;

Query Match 100.0%; Score 21; DB 1; Length 227;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 Db 199 SYDA 202

## RESULT 29

Y1B8\_AGR9 STANDARD; PRT; 230 AA.  
 ID Y1B8\_AGR9  
 AC P05357;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Y1B8 protein.  
 GN Y1B8.  
 OS Agrobacterium tumefaciens (strain 15955).  
 OG Plasmid pTi15955.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Rhizobium.  
 OX NCBI\_TaxID=190386;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=88247765; PubMed=2837739;  
 RA Thompson D.V., Melchers L.S., Idler K.B., Shilpercoort R.A.,  
 RA Hooykaas P.J.J.;  
 RT "Analysis of the complete nucleotide sequence of the Agrobacterium  
 RT tumefaciens Y1B8 operon.";  
 RT Nucleic Acids Res. 16:4621-4636(1988).  
 RL Nucleic Acids Res. 16:4621-4636(1988).  
 CC -I- FUNCTION: VIRB PROTEINS ARE SUGGESTED TO ACT AT THE BACTERIAL  
 CC SURFACE AND THERE PLAY AN IMPORTANT ROLE IN DIRECTING T-DNA  
 CC TRANSFER TO PLANT CELLS.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X06826; CAC15165.1; AYT\_INIT.  
 DR PIR: S00784; B8AG55.

```

KM Crown gall tumor; Plasmid.
SQ SEQUENCE 230 AA; 25408 MW; 92D000C464810D6D CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 230;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
Db 104 SYDA 107

RESULT 30
VIRB_AGRU STANDARD; PRT; 230 AA.
ID VIRB_AGRU
AC P09781;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE VIRB8 protein.
GN Agrobacterium tumefaciens.
OS Plasmid pTiA6.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
CC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=358;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88186901; PubMed=3281947;
RA Ward J.E., Akiyoshi D.E., Regier D., Datla A., Gordon M.P.,
RA Nester E.W.;
RT "Characterization of the virB operon from an Agrobacterium
RL J. Biol. Chem. 263:5804-5814(1988).
RN [2]
RP REVISIONS.
RX MEDLINE=90170994; PubMed=2307685;
RA Ward J.E., Akiyoshi D.E., Regier D., Datla A., Gordon M.P.,
RA Nester E.W.;
RL J. Biol. Chem. 265:4768-4768(1990).
CC -1- FUNCTION: VIRB PROTEINS ARE SUGGESTED TO ACT AT THE BACTERIAL
CC SURFACE AND THERE PLAY AN IMPORTANT ROLE IN DIRECTING T-DNA
CC TRANSFER TO PLANT CELLS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J03216; -; NOT_ANNOTATED_CDS.
DR PIR; I28621; B8AGA6.
DR Crown gall tumor; Plasmid.
SQ SEQUENCE 230 AA; 25382 MW; 87CB1A7A8B30BD76 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 230;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
Db 104 SYDA 107

RESULT 31
Y416_CHLTR STANDARD; PRT; 236 AA.
AC 084421;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

```

```

DE Probable metal transport system ATP-binding protein CTA16.
GN CTA16.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/1W-3/CX;
RX MEDLINE=9000809; PubMed=9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis."
RL Science 282:754-759(1998).
CC -1- FUNCTION: PART OF AN ATP-DRIVEN TRANSPORT SYSTEM CTA15/CTA16/CTA17
CC FOR A METAL. PROBABLY RESPONSIBLE FOR ENERGY COUPLING TO THE
CC TRANSPORT SYSTEM.
CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC -1- SIMILARITY: BELONGS TO THE ABC_TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001315; AAC68013.1; -
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transport.
DR Pfam; PF00005; ABC_tran.1.
DR SMART; SM00382; AAA.1.
DR PROSITE; PS00211; ABC_TRANSPORTER.1.
RX Hypothetical protein; Transport; Inner membrane; ATP-binding;
KW Complete proteome.
FT NP_BIND 39 46 ATP (POTENTIAL).
SQ SEQUENCE 236 AA; 26193 MW; DE74774507950365 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 236;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
Db 84 SYDA 87

RESULT 32
VIRB_AGRU5 STANDARD; PRT; 237 AA.
ID VIRB_AGRU5
AC P17758;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE VIRB8 protein.
GN VIRB8 OR ATU6174 OR AGR_PTI.11.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Plasmid pTiC58.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90318324; PubMed=2370849;
RA Kuldau G.A., de Vos G., Owen J., McCaffrey G., Zambryski P.;
RT "The virB operon of Agrobacterium tumefaciens pTiC58 encodes 11 open
RT reading frames."
RL Mol. Gen. Genet. 221:256-266(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90301800; PubMed=2194232;

```

```

RA Rogovsky P.M., Powell B.S., Shirasu K., Lin T.-S., Morel P.,
RA Zyprian E.M., Steck T.R., Kado C.I.,
RT Molecular characterization of the vtr regulon of Agrobacterium
RT tumefaciens: complete nucleotide sequence and gene organization of
RT the 28.63-kbp regulon cloned as a single unit.
RL Plasmid 23:85-106(1990).
RN (3)
RP SEQUENCE FROM N.A.
RX MEDLINE-21608550; PubMed-11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitaajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayvin T., Levy R., Li M.-J., McCalland E., Palmeri A., Gordon D.,
RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nestor E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RL Science 294:2317-2323(2001).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE-21608551; PubMed-11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Ouriolo B., Goldman B.S., Cao Y., Askenazi M., Hailing C., Mullin L.,
RA Houmlet K., Gordon J., Vaudin M., Iarchenou O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughly D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RL Science 294:2323-2328(2001).
RN -1- FUNCTION: VIRB PROTEINS ARE SUGGESTED TO ACT AT THE BACTERIAL
RN SURFACE AND THERE PLAY AN IMPORTANT ROLE IN DIRECTING T-DNA
RN TRANSFER TO PLANT CELLS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC CC
DR EMBL: X53264; CA37361.1; -
DR EMBL: J03320; AAA91598.1; -
DR EMBL: AE009436; AAL46410.1; -
DR EMBL: AE007923; AAK90936.1; ALT_INIT.
DR PIR: S12348; B8AG58.
KM Crown gall tumor; Plasmid; Complete proteome.
FT CONFLICT 21 22 MISSING (IN REF. 1).
FT CONFLICT 128 129 SA -> TR (IN REF. 2).
SQ SEQUENCE 237 AA; 26294 MW; 71FEF1084B9597E9 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 237;
Best Local Similarity 100.0%; Pred. NO. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4
DB 111 SYDA 114

RESULT 33
PCRB_SCHGR STANDARD; PRT; 250 AA.
AC PE2886;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative beta-carotene binding protein.

```

```

OS Schistocerca gregaria (Desert locust).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pleurogata; Neoptera; Orthopteroidea; Orthoptera; Caellifera;
OC Acridocephala; Acridoidea; Acrididae; Cyrtacanthacridinae;
OC Schistocerca.
OX NCBI_TaxId=7010;
RN [1]
RP SEQUENCE.
RC TISSUE=Cuticle;
RX MEDLINE-21468012; PubMed-11583931;
RA Wybrandt G.B., Andersen S.O.;
RT "Purification and sequence determination of a yellow protein from
RT sexually mature males of the desert locust, Schistocerca gregaria."
RL Insect Biochem. Mol. Biol. 31:1183-1189(2001).
CC -1- FUNCTION: Has beta-carotene-binding activity. May be involved in
CC the transport of carotenes from internal tissues to epidermis and
CC cuticle of the locust.
CC -1- TISSUE SPECIFICITY: Deposited in the epidermis and cuticle of male
CC locusts during their sexual maturation.
DR InterPro: IPR004272; Odorant-binding.
DR Pfam: PF03027; DUF233; 1.
KW Transport.
SQ SEQUENCE 250 AA; 25666 MW; E6074CF88D822EB5 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 250;
Best Local Similarity 100.0%; Pred. NO. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4
DB 215 SYDA 218

RESULT 34
Y625_TREPA
ID Y625_TREPA STANDARD; PRT; 250 AA.
AC O83633;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TP0625.
GN TP0625.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxId=160;
RN [1]
RP SEQUENCE FROM N.A.
RN STRAIN=Nichols;
RX MEDLINE-98332770; PubMed-9665876;
RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,
RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richard D., Howell J.K., Chidambaram M., Utterback T.,
RA McDonald L., Artlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete."
RL Science 281:375-388(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC CC
DR EMBL: AE001237; AAC65604.1; -
DR TIGR: TP0625; -
KW Hypothetical protein; Coiled coil; Complete proteome.
FT DOMAIN 165 208 COILED COIL (POTENTIAL).
SQ SEQUENCE 250 AA; 28874 MW; 107002FF47B25186 CRC64;

```

Query Match 100.0%; Score 21; DB 1; Length 250;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 DB 81 SYDA 84

## RESULT 35

RL7C\_SCHPO STANDARD; PRT; 251 AA.  
 ID RL7C\_SCHPO  
 AC 060143;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 60S ribosomal protein L7-C.  
 GN RPL7C OR SPBC18H10.12C.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 NCBI\_TaxID=4896;  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA McManus P., Mouton S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Stevens K., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volkart G., Aert R., Robben J., Gymnopoulos B.,  
 RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Rhode G.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe."  
 RL Nature 415:871-880(2002).  
 CC -1- MISCELLANEOUS: THERE ARE THREE GENES FOR L7 IN S.POMBE.  
 CC -1- SIMILARITY: BELONGS TO THE L30P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AL022304; CAA18409.1; -  
 CC InterPro: IPR000517; Ribosomal\_L30.  
 CC Pfam: PF00327; Ribosomal\_L30; 1.  
 CC TrIRFAMS: TIGR01310; L7; 1.  
 CC PROSITE: PS00634; RIBOSOMAL\_L30; 1.  
 KW Ribosomal protein; Multigene family;  
 SQ SEQUENCE 251 AA; 28730 MW; 5382A27E39F3800E CRC64;

Query Match 100.0%; Score 21; DB 1; Length 251;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 DB 58 SYDA 61

## RESULT 36

SOI\_TREPA STANDARD; PRT; 253 AA.  
 ID SOI\_TREPA  
 AC 083296;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Protein soj homolog.  
 GN TP0272.  
 OS Treponema pallidum.  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.  
 OX NCBI\_TaxID=160;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NICHOLS;  
 RX MEDLINE=98332770; PubMed=9665876;  
 RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,  
 RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,  
 RA Koderger E., Hardham J.M., McLeod M.P., Saldberg S., Peterson J.,  
 RA Khakh H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,  
 RA McDonald L., Artlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,  
 RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,  
 RA Venter J.C.;  
 RT "Complete genome sequence of Treponema pallidum, the syphilis  
 RT spirochete."  
 RL Science 281:375-388(1998).  
 CC -1- SIMILARITY: BELONGS TO THE PARA FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AE001208; AAC65260.1; -  
 CC TIGR: TP0272; -  
 DR InterPro: IPR000707; ATPase\_Para.  
 DR Pfam: PF00991; Para; 1.  
 KW ATP-binding; Complete proteome.  
 FT NP\_BIND 10 17 ATP (POTENTIAL).  
 SQ SEQUENCE 253 AA; 27341 MW; 284B3B31A9306EDA CRC64;  
 Query Match 100.0%; Score 21; DB 1; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## RESULT 37

BIOC\_SERMA STANDARD; PRT; 255 AA.  
 ID BIOC\_SERMA  
 AC P36571;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Biotin synthesis protein bioc.  
 GN BIOC.  
 OS Serratia marcescens.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;



OC Serratia.  
 RX NCB1\_TaxID=615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sc41;  
 RX MEDLINE=94071435; PubMed=8250549;  
 RA Sakurai N., Imai Y., Masuda M., Komatsubara S., Tosa T.;  
 RT "Molecular breeding of a biotin-hyperproducing *Serratia marcescens*  
 str. Sc41".  
 RL Appl. Environ. Microbiol. 59:3225-3232(1993).  
 CC -1- FUNCTION: BIOC IS INVOLVED IN AN EARLY, BUT CHEMICALLY UNEXPLORED,  
 STEP IN THE CONVERSION OF PIMELIC ACID TO BIOTIN.  
 CC -1- PATHWAY: Biotin biosynthesis.  
 CC -1- SIMILARITY: BELONGS TO THE METHYLTRANSFERASE SUPERFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: D17468; BAA04287.1; -;  
 DR InterPro: IPR000051; SAM\_bind.  
 DR InterPro: IPR004033; UbiE/COO5\_Metrf.  
 DR Pfam: PF01209; UbiE\_methyltran; 1.  
 KW Biotin biosynthesis; Transferase; Methyltransferase.  
 SQ SEQUENCE 255 AA; 27776 MW; 1DD2532B3FCA7F67 CRC64;  
 Query Match 100.0%; Score 21; DB 1; Length 255;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 SYDA 4  
 Db 23 SYDA 26  
 RESULT 38  
 CAH2\_CHICK STANDARD; PRT; 259 AA.  
 ID CAH2\_CHICK  
 AC P07630;  
 DT 01-APR-1988 (Rel. 07, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Carbonic anhydrase II (EC 4.2.1.1) (Carbonate dehydratase II) (CA-II).  
 GN CA2.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCB1\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Hubbard White Mountain; TISSUE=Testis;  
 RX MEDLINE=95011620; PubMed=7926806;  
 RA Mezquita J., Pau M., Mezquita C.;  
 RT "A novel carbonic anhydrase II mRNA isolated from mature chicken  
 RT testis displays a TATA box and other promoter sequences in a leader  
 RT 5' untranslated region not present in somatic tissues".  
 RL Gene 147:231-235(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=White Leghorn;  
 RX MEDLINE=87146391; PubMed=3029691;  
 RA Yoshihara C.M., Lee J.-D., Dodgson J.B.;  
 RT "The chicken carbonic anhydrase II gene: evidence for a recent shift  
 RT in intron position".  
 RL Nucleic Acids Res. 15:753-770(1987).  
 RN [3]  
 RP SEQUENCE OF 4-259 FROM N.A.  
 RC STRAIN=White Leghorn; TISSUE=Retina;

RX MEDLINE=87133522; PubMed=3102231;  
 RA Rogers J.H.;  
 RT "Sequence of carbonic anhydrase II cDNA from chick retina".  
 RL Eur. J. Biochem. 162:119-122(1987).  
 RN [4]  
 RP SEQUENCE OF 221-259 FROM N.A.  
 RC TISSUE=Retina;  
 RX MEDLINE=90192090; PubMed=1969140;  
 RA Godbout R., Andison R., Upton C., Day R.;  
 RT "Utilization of the second polyadenylation signal at the 3' end of  
 RT the chicken carbonic anhydrase II gene".  
 RL Nucleic Acids Res. 18:1049-1049(1990).  
 RN [5]  
 RP SEQUENCE OF 7-86 FROM N.A.  
 RX MEDLINE=84255154; PubMed=6331256;  
 RA Yoshihara C.M., Federspiel M., Dodgson J.B.;  
 RT "Isolation of the chicken carbonic anhydrase II gene".  
 RL Ann. N.Y. Acad. Sci. 429:332-334(1984).  
 CC -1- FUNCTION: REVERSIBLE HYDRATATION OF CARBON DIOXIDE.  
 CC -1- CATALYTIC ACTIVITY: H(2)CO(3) -> CO(2) + H(2)O.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC-TYPE CARBONIC ANHYDRASE  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Z14957; CAA78681.1; -;  
 DR EMBL: X12639; CAA31175.1; -;  
 DR EMBL: X06000; CAA29417.1; -;  
 DR EMBL: X06001; CAA29417.1; JOINED.  
 DR EMBL: X06002; CAA29417.1; JOINED.  
 DR EMBL: X06003; CAA29417.1; JOINED.  
 DR EMBL: X06004; CAA29417.1; JOINED.  
 DR EMBL: X06005; CAA29417.1; JOINED.  
 DR EMBL: X04810; CAA28501.1; -;  
 DR EMBL: X17378; CAA3250.1; -;  
 DR EMBL: M25943; AAA48646.1; -;  
 DR PIR: S01078; S01078.  
 DR PIR: A26415; A26415.  
 DR PIR: S10229; S10229.  
 DR PIR: S31987; S31987.  
 DR HSSP: P00918; ICIM.  
 DR InterPro: IPR001148; Euk\_Coanhd.  
 DR Pfam: PF00194; carb\_anhydrase; 1.  
 DR ProDom: PD000865; Euk\_Coanhd; 1.  
 DR PROSITE: PS00162; Euk\_CO2\_ANHYDRASE; 1.  
 KW lyase; Zinc.  
 FT INIT MET 0  
 FT METAL 93 0  
 FT METAL 95 95 ZINC (CATALYTIC).  
 FT METAL 118 118 ZINC (CATALYTIC).  
 FT CONFLICT 4 4 W -> L (IN REF. 3).  
 FT CONFLICT 7 7 D -> G (IN REF. 5).  
 FT CONFLICT 86 86 V -> S (IN REF. 5).  
 FT CONFLICT 249 249 L -> V (IN REF. 2).  
 SQ SEQUENCE 259 AA; 28876 MW; DF133083864A4569 CRC64;  
 Query Match 100.0%; Score 21; DB 1; Length 259;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 SYDA 4  
 Db 49 SYDA 52  
 RESULT 39

EXL2\_ARATH STANDARD: PRT: 265 AA.

AC Q95VE5:

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DE Expansin-like 2 precursor (At-EXPL2) (Ath-ExpBeta-2.2).

GN EXPL2 OR AT4G38400 OR F22I13.170.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OC NCBI\_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;

RX MEDLINE=20083488; PubMed=10617198;

RA Meyer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G., Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terry N., Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M., Weichselgartner M., de Simone V., Obermaier B., Macho R., Mueller M., Kreis M., Delseny M., Puigdomenech P., Watson M., Schlottheim T., Reichert B., Portelelle D., Perez-Alonso M., Boutry M., Bancroft I., Vos P., Hohelsel J., Zimmermann W., Wedler H., Ridley P., Langham S.-A., McCullagh B., Bilham L., Robben J., Van der Schueren J., Gymonpretz B., Chuang Y.-d., Vandenbussche F., Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E., Wiltzengerger T., Bothe G., Ramsperger U., Hilbert H., Braun M., Holzner E., Brandt A., Peters S., van Staveren M., Dirkse W., Moolman P., Klein Lankhorst R., Rose M., Hauf J., Koetler P., Bernsen S., Hempel S., Feldpausch M., Lambrecht S., Van den Daele H., de Keyser A., Bysschaert C., Gielen J., Villarroel R., De Clercq R., Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S., Clark L., Doggett J., Hall S., Kay M., Lennard N., McKay K., Mayes R., Borkova D., Bloeker H., Scharfe M., Grimm M., Loehnert T.-H., Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Partmann B., Granderath K., Dauner D., Herzl A., Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E., Massenet O., Quigley F., Clabaud G., Muendlein A., Felber R., Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S., Chefor F., Cooke R., Berger C., Monfort A., Casacuberta E., Gibbons T., Weber N., Berger C., Monfort A., Casacuberta E., Perez-Perez A., Punnelle B., Bent E., Johnson S., Terol J., Torres A., Hejnen D., Haase D., Lemcke K., Mewes H.-W., Stocker S., Frisman D., Bevan M., Wilson R.K., de la Bastide M., Habermann K., Zaccaria P., Dehla N., Gnoj L., Schutz K., Huang E., Spiegel L., Sektion M., Murray J., Sheet P., Cordes M., Abu-Threideh J., Stoneking T., Kallicki J., Graves T., Harmon G., Edwards J., Latelle P., Courtney L., Clout J., Abbott A., Scott K., Johnson D., Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K., Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L., Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D., Du H., Ali J., Berghoff A., Jones K., Drome K., Cotton M., Joshi C., Antonoli B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C., Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R., Swamy I.K., O'Shaughnessy A., Rodriguez X., Hoffman J., Tili S., Granat S., Shohdy N., Hasegawa A., Hameed A., Lochi M., Johnson A., Chen E., Maria M., Martensen R., McCombie W.R.;

RA "Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana."

RT Nature 402:769-777(1999).

RL [2]

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;

RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;

RT "RIKEN Arabidopsis full length cDNA clones (RAFLS) sequenced by the SSP consortium (Salk/Stanford/GREC)."

CC Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO THE EXPANSIN FAMILY.

CC -1- SIMILARITY: BELONGS TO THE EXPANSIN-LIKE EG45 DOMAIN.

CC -1- SIMILARITY: BELONGS TO THE EXPANSIN-LIKE CBD DOMAIN.

CC -1- SIMILARITY: BELONGS TO THE EXPANSIN-LIKE CBD DOMAIN.

CC -1- DATABASE: NAME=EXPANSIN homepage;  
WWW="http://www.bio.psu.edu/expansin/".

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-stb.ch/announce/>  
CC or send an email to [license@isb-stb.ch](mailto:license@isb-stb.ch)).

CC -----

DR EMBL: AL035539; CAB37486.1; -

DR EMBL: AL161593; CAB80505.1; -

DR EMBL: AY050470; AAK91483.1; -

DR EMBL: AF378855; AAK55658.1; -

DR HSSP: P43214; IWHO.

DR InterPro: IPR000882; Pollen.allergen.1.

DR Pfam: PF01357; Pollen.allergen.1.

DR PRINTS: PR01225; EXPANSINFAMILY.

DR PRODOM: PD02179; Pollen.allergen.1.

DR PROSITE: PS50843; EXPANSIN.CBD.1.

DR PROSITE: PS50842; EXPANSIN.EG45.1.

KM Glycoprotein; Signal; Multigene family.

FT SIGNAL 1 21

FT CHAIN 22 265

FT DOMAIN 42 148

FT DOMAIN 162 244

FT CARBOHYD 100 100

FT CARBOHYD 103 103

FT SQUENCE 265 AA; B57BAC921899F40D CAC64;

Query Match Score: 21; DB 1; Length 265;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4

DB 241 SYDA 244

RESULT 40

FLIP\_CAUCR STANDARD: PRT: 266 AA.

ID Q45980;

AC 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DE Flagellar biosynthetic protein Flp.

GN FLIP OR CC0951.

OS Caulobacter crescentus.

OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;

OC Caulobacter.

OC NCBI\_TaxID=155892;

OX [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 19089 / CB15;

RX MEDLINE=95325304; PubMed=7601828;

RA Guber J.W., Boyd C.H., Jarvis M., Mangan E.K., Rizzo M.F., Windrova J.A.;

RT "Temporal and spatial regulation of flp, an early flagellar gene of Caulobacter crescentus that is required for motility and normal cell division."

RT J. Bacteriol. 177:3656-3667(1995).

RL [2]

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 19089 / CB15;

RX MEDLINE=21173698; PubMed=11259647;

RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E., Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R., Picocha I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B., DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H., Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K., Ullrich T., Tian K., Wolf A., Yamatyan J., Ermolaeva M., White O.,

RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of *Caulobacter crescentus*.";   
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 CC -1- FUNCTION: MAY BE A COMPONENT OF THE FLAGELLUM. IT IS REQUIRED FOR  
 CC NORMAL CELL DIVISION. MAY BE IMPLICATED IN THE SECRETION OF  
 CC VIRULANCE FACTORS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE FLIP/MOPC/SPAP FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U20387; AAA66882.1; -  
 DR EMBL: AE005772; AAK22935.1; -  
 DR TIGR: CC0951; -  
 DR InterPro: IPR002039; FLIP.  
 DR Pfam: PF00813; FLIP; 1.  
 DR ProDom: PD002586; FLIP; 1.  
 DR TIGRFAMs: TIGR01103; FLIP; 1.  
 DR PROSITE: PS01060; FLIP\_1; 1.  
 DR PROSITE: PS01061; FLIP\_2; 1.  
 KM Flagella; Transmembrane; Complete proteome.  
 FT TRANSMEM 20 40 POTENTIAL.  
 FT TRANSMEM 58 78 POTENTIAL.  
 FT TRANSMEM 102 122 POTENTIAL.  
 FT TRANSMEM 202 222 POTENTIAL.  
 FT TRANSMEM 226 246 POTENTIAL.  
 SO SEQUENCE 266 AA; 28527 MW; A84F17CB1C65A947 CRC64;  
 Query Match 100.0%; Score 21; DB 1; Length 266;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SYDA 4  
 Db 125 SYDA 128  
 RESULT 41  
 PANB\_CORGL STANDARD; PRT; 271 AA.  
 ID PANB\_CORGL  
 AC Q9X712;  
 DT 16-OCT-2001 (Rel. 40; Created)  
 DT 16-OCT-2001 (Rel. 40; Last sequence update)  
 DT 15-JUN-2002 (Rel. 41; Last annotation update)  
 DE 3-methyl-2-oxobutanate hydroxymethyltransferase (EC 2.1.2.11)  
 DE (ketopantoate hydroxymethyltransferase).  
 GN PANB OR CGL0114.  
 OS Corynebacterium glutamicum (Brevibacterium flavum).  
 CC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;  
 CC Actinomycetales; Corynebacteriaceae; Corynebacteriaceae;  
 CC Corynebacterium.  
 CC NCBI\_TaxID=1718;  
 RN NCBI [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 13032 / DSM 20300 / NCIB 10025;  
 RX MEDLINE=99240418; PubMed=10232988;  
 RA Eggeling L., Sahm H.;  
 RT "D-Pantothenate synthesis in *Corynebacterium glutamicum* and use of  
 RT panB and genes encoding L-valine synthesis for D-pantothenate  
 RT overproduction.";   
 RL Appl. Environ. Microbiol. 65:1973-1979(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 13032 / DSM 20300 / NCIB 10025;  
 RA Nakagawa S.;  
 RT "Complete genomic sequence of *Corynebacterium glutamicum* ATCC 13032.";   
 RT Submitted (May-2002) to the EMBL/GenBank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: 5,10-methylenetetrahydrofolate + 3-methyl-2-  
 CC oxobutanate = tetrahydrofolate + 2-dehydropanoate.  
 CC -1- PATHWAY: Pantothenate biosynthesis; first branch; first step.  
 CC -1- SIMILARITY: BELONGS TO THE PANB FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X96580; CAA65397.1; -  
 DR EMBL: AP005274; BAB97507.1; -  
 DR InterPro: IPR003700; Pantoate\_transf.  
 DR Pfam: PF02548; Pantoate\_transf; 1.  
 DR TIGRFAMs: TIGR00222; panb; 1.  
 KM Pantothenate biosynthesis; Transferase; Methyltransferase.  
 SO SEQUENCE 271 AA; 28552 MW; 0D5A8DCA93A03FA CRC64;  
 Query Match 100.0%; Score 21; DB 1; Length 271;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SYDA 4  
 Db 31 SYDA 34  
 RESULT 42  
 CYNL\_SYNY3 STANDARD; PRT; 274 AA.  
 ID CYNL\_SYNY3  
 AC Q54735; P74088;  
 DT 01-NOV-1997 (Rel. 35; Created)  
 DT 01-NOV-1997 (Rel. 35; Last sequence update)  
 DT 16-OCT-2001 (Rel. 40; Last annotation update)  
 DE Carbonic anhydrase (EC 4.2.1.1).  
 DE ICEA OR SLR1347.  
 GN Synecocystis sp. (strain PCC 6803).  
 OS Bacteria; Cyanobacteria; Chroococcales; Synecocystis.  
 CC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.  
 CC NCBI\_TaxID=1148;  
 RN NCBI [1]  
 RP SEQUENCE FROM N.A.  
 RA So A.K.C., Espie G.S.;  
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97061201; PubMed=8905231;  
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
 RA Miyajima N., Hirosewa M., Sugita M., Sasamoto S., Kimura T.,  
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,  
 RA Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,  
 RA Yamada M., Yasuda M., Tabata S.;  
 RT "Sequence analysis of the genome of the unicellular cyanobacterium  
 RT Synecocystis sp. strain PCC6803. II. Sequence determination of the  
 RT entire genome and assignment of potential protein-coding regions.";   
 RL DNA Res. 3:109-136(1996).  
 CC -1- FUNCTION: REVERSIBLE HYDRATATION OF CARBON DIOXIDE. ESSENTIAL  
 CC TO PHOTOSYNTHETIC CARBON DIOXIDE FIXATION.  
 CC -1- CATALYTIC ACTIVITY: H(2)CO(3) = CO(2) + H(2)O.  
 CC -1- SIMILARITY: BELONGS TO THE PLANT AND PROKARYOTIC CARBONIC  
 CC ANHYDRASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U45962; AAC46375.1; -

DR EMBL; D90912; BAA18166.1; ALT\_INIT.  
 DR InterPro; IPR001765; PROK\_COAHnd.  
 DR Pfam; PF00484; PRO\_CA; 1.  
 DR PROSITE; PS00704; PROK\_CO2\_ANNHYDRASE\_1; 1.  
 DR PROSITE; PS00705; PROK\_CO2\_ANNHYDRASE\_2; 1.  
 DR Lyase; Zinc; Complete proteome.  
 KW CONFLICT 219 219 P -> L (IN REF. 2).  
 FT SEQUENCE 274 AA; 30761 MW; 2FAF634246C0D5A9 CRC64;  
 SQ  
 Query Match 100.0%; Score 21; DB 1; Length 274;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 Db 227 SYDA 230

RESULT 43  
 YSIL\_CAEEL STANDARD; PRT; 279 AA.  
 ID YSIL\_CAEEL  
 AC Q10033; Q903J1;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical protein F1569.1 in chromosome X.  
 GN F1569.1  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Bristol N2;  
 RA Sulston J.E.;  
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP REVISIONS, AND ALTERNATIVE SPLICING.  
 RA Durbin R.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a (shown here) and b; may be  
 CC produced by alternative splicing.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; 247068; CAA87331.1; -  
 DR EMBL; 247068; CAB54215.1; -  
 DR WormPep; F1569.1a; CE01549.  
 DR WormPep; F1569.1b; CE23660.  
 KW Hypothetical protein; Alternative splicing.  
 FT DOMAIN 2 64 ARG/SER-RICH.  
 FT DOMAIN 87 150 GLU-RICH.  
 FT DOMAIN 142 150 POLY-GLU.  
 FT VARSPLIC 1 74 MISSING (IN ISOFORM B).  
 SQ SEQUENCE 279 AA; 30723 MW; D196C358F754BEA6 CRC64;  
 Query Match 100.0%; Score 21; DB 1; Length 279;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 Db 92 SYDA 95

RESULT 44  
 SUBB\_METTH

ID SUBB\_METTH STANDARD; PRT; 280 AA.  
 AC O26957;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Inositol-1-monophosphatase (EC 3.1.3.25) (IMPase) (Inositol-1-  
 DE phosphatase) (I-1-Pase).  
 GN SUBB OR MTH871.  
 OS Methanobacterium thermoautotrophicum.  
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;  
 OC Methanobacteriaceae; Methanothermobacter.  
 OX NCBI\_TaxID=187420;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Delta H;  
 RX MEDLINE=98037514; PubMed=9371463;  
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,  
 RA Aldredge T., Bashirzadeh R., Blakey D., Cook R., Gilbert K.,  
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,  
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,  
 RA Jivani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,  
 RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,  
 RA Daniels C.J., Mao J.-I., Rice P., Nollig J., Reeve J.N.;  
 RT deltaH: functional analysis and comparative genomics."  
 RL J. Bacteriol. 179:7135-7155(1997).  
 CC -1- CATALYTIC ACTIVITY: Myo-Inositol 1-monophosphate + H(2)O = myo-  
 CC inositol + phosphate.  
 CC  
 CC -1- COPFACTOR: Magnesium (By similarity).  
 CC  
 CC -1- SIMILARITY: BELONGS TO THE INOSITOL MONOPHOSPHATASE FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AE000863; AAB85369.1; -  
 DR HSSP; P46275; IDB2.  
 DR InterPro; IPR000760; Inositol\_P.  
 DR Pfam; PF00459; Inositol\_P; 1.  
 DR PROSITE; PS00629; IMP\_2; 1.  
 DR PROSITE; PS00630; IMP\_2; 1.  
 KW Hydrolyase; Magnesium; Complete proteome.  
 SQ SEQUENCE 280 AA; 30173 MW; FOAE435970AED882 CRC64;  
 Query Match 100.0%; Score 21; DB 1; Length 280;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 Db 211 SYDA 214

RESULT 45  
 DEGV\_BACSU STANDARD; PRT; 281 AA.  
 ID DEGV\_BACSU  
 AC P32436;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE DegV protein.  
 GN DEGV.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=94018599; PubMed=8412657;

RA Londono-Vallejo J.A., Dubnau D.;  
 RT "comp. a Bacillus subtilis late competence locus, encodes a protein  
 RT similar to ATP-dependent RNA/DNA helicases.";

RA Mol. Microbiol. 9:119-131(1993).

RA [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RA Soldo B., Lazarevic V., Manuel C., Karamata D.;  
 RT "Sequence of the Bacillus subtilis 168 chromosomal region from 305  
 RT to 307 degree.";  
 RT Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.

RA [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RA MEDLINE=98044033; PubMed=9384377;  
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
 RA Azevedo V., Bertolo M.G., Bessieres P., Boilot A., Borchert S.,  
 RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,  
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
 RA Ertlan K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,  
 RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,  
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
 RA Joris B., Karamata D., Kasahara Y., Klier-Blanchard M., Klein C.,  
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
 RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,  
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
 RA Parro V., Pohl T.M., Portetalle D., Portolillo S., Prescott A.M.,  
 RA Prescan E., Puig C., Punelle B., Rapoport G., Rey M., Reynolds S.,  
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,  
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
 RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,  
 RA Sorokin A., Taccioni E., Takagi T., Takahashi H., Takemaru K.,  
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpestra P., Tognoni A.,  
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
 RA Viart A., Wambutt R., Wedler E., Wedler H., Weitzenecker K., Yata K.,  
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;  
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus  
 RT subtilis";  
 RT Nature 390:249-256(1997).

RA [4]  
 RP SEQUENCE OF 1-107 FROM N.A.  
 RX MEDLINE=8903891; PubMed=3141378;  
 RA Henner D.J., Yang M., Ferrari E.;  
 RT "Localization of Bacillus subtilis sacU(hy) mutations to two linked  
 RT genes with similarities to the conserved procaryotic family of two-  
 RT component signalling systems.";  
 RT J. Bacteriol. 170:5102-5109(1988).

RA -1- SIMILARITY: TO S.AUREUS PLASMID PSK1 HYPOTHETICAL 15.5 KDA PROTEIN  
 CC (AC P13977) AND TO M.GENTILICUM MG326.

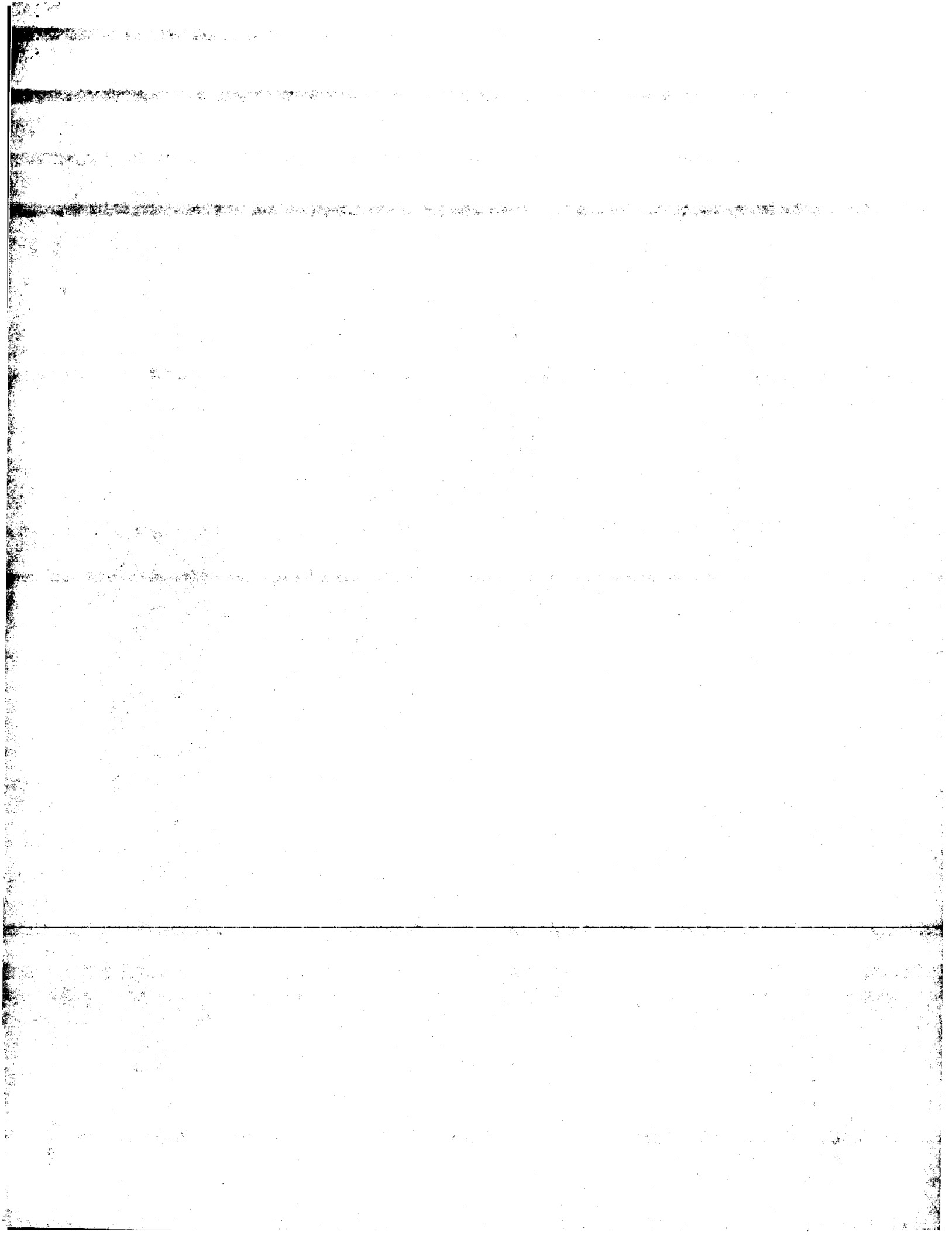
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL: Z18629; CAA79225.1; -;  
 CC EMBL: U56901; AAC44939.1; -;  
 CC EMBL: Z99122; CAB15565.1; -;  
 CC EMBL: M23558; -; NOT\_ANNOTATED\_CDS.  
 CC PIR: D30191; D30191.  
 CC PIR: S28596; S28596.  
 CC Subtilist; BG10394; degV.  
 CC Interpro; IPR003797; degV.

DR Pfam: PF02645; DUF194; 1.  
 DR TIGRfams: TIGR00762; degV; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 281 AA; 31580 MM; 737B6DC65D6FCE64 CRC64;  
 Query Match 100.0%; Score 21; DB 1; Length 281;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
 ||||  
 Db 79 SYDA 82

Search completed: February 6, 2003, 11:16:46  
 Job time : 7.33333 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 11:16:00 : Search time 21.3333 Seconds  
(without alignments)  
38.634 Million cell updates/sec

Title: PAT943-3  
Perfect score: 21  
Sequence: 1 syda 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapect 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 1226

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SPREMBL\_21:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mammal:\*  
8: sp\_mhc:\*  
9: sp\_organelle:\*  
10: sp\_plant:\*  
11: sp\_protist:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	21	100.0	40	5 Q9NFH5	Q9NFH5 plasmodium
2	21	100.0	53	16 Q8YX6	Q8YX6 anabaena sp
3	21	100.0	56	3 Q12081	Q12081 saccharomyc
4	21	100.0	67	2 Q9X374	Q9X374 bacillus an
5	21	100.0	75	16 Q9RIC5	Q9RIC5 yersinia an
6	21	100.0	77	9 Q9MCA1	Q9MCA1 bacterioph
7	21	100.0	78	2 Q9Z3S8	Q9Z3S8 pseudomonas
8	21	100.0	80	16 Q9Z7E0	Q9Z7E0 listeria in
9	21	100.0	81	2 Q9ZGZ2	Q9ZGZ2 yersinia pe
10	21	100.0	81	2 Q9K510	Q9K510 pseudomonas
11	21	100.0	81	16 Q8Y3X8	Q8Y3X8 listeria mo
12	21	100.0	82	5 Q9VOT7	Q9VOT7 drosophila
13	21	100.0	84	16 Q8Z021	Q8Z021 salmoneilla
14	21	100.0	84	16 Q8X8N9	Q8X8N9 escherichia
15	21	100.0	85	2 Q68916	Q68916 streptomyces
16	21	100.0	85	16 Q8Z7L4	Q8Z7L4 salmoneilla

17	21	100.0	87	5 Q9GY45	Q9GY45 leishmania
18	21	100.0	92	12 Q72135	Q72135 human papil
19	21	100.0	93	11 Q91Z41	Q91Z41 mus musculu
20	21	100.0	95	2 Q85877	Q85877 sphingomona
21	21	100.0	96	16 Q8Y0B6	Q8Y0B6 ralsomnia s
22	21	100.0	99	3 Q14371	Q14371 schizosacch
23	21	100.0	100	10 Q8S0G9	Q8S0G9 oryza sativ
24	21	100.0	100	16 Q8RH72	Q8RH72 fusobacteri
25	21	100.0	103	2 Q8S916	Q8S916 sphingomona
26	21	100.0	106	16 Q9A108	Q9A108 streptococ
27	21	100.0	106	17 Q9HN89	Q9HN89 halobacteri
28	21	100.0	109	10 Q8W5F8	Q8W5F8 oryza sativ
29	21	100.0	111	10 Q94ZU5	Q94ZU5 oryza sativ
30	21	100.0	112	5 Q9NBJ7	Q9NBJ7 caenorhabdi
31	21	100.0	113	2 Q9XCB3	Q9XCB3 rhodothermu
32	21	100.0	113	5 Q95WK3	Q95WK3 anopheles p
33	21	100.0	113	16 Q9RZP8	Q9RZP8 delinococcus
34	21	100.0	115	2 Q8RSN6	Q8RSN6 uncultured
35	21	100.0	115	16 Q9ZJ33	Q9ZJ33 rickettsia
36	21	100.0	116	14 Q991X7	Q991X7 uncultured
37	21	100.0	121	16 Q9ZUS4	Q9ZUS4 rhizobium m
38	21	100.0	127	5 Q9XVLO	Q9XVLO caenorhabdi
39	21	100.0	129	16 P73361	P73361 synchocyst
40	21	100.0	130	16 Q99WV3	Q99WV3 staphylococ
41	21	100.0	132	5 Q9NBR3	Q9NBR3 caenorhabdi
42	21	100.0	132	16 Q8X858	Q8X858 escherichia
43	21	100.0	133	5 Q9XUK1	Q9XUK1 caenorhabdi
44	21	100.0	133	16 Q97HT8	Q97HT8 clostridium
45	21	100.0	133	16 Q97FH2	Q97FH2 clostridium

## ALIGNMENTS

## RESULT 1

ID Q9NFH5 PRELIMINARY; PRT; 40 AA.  
AC Q9NFH5;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE Hypothetical 4.8 kDa protein (Fragment).  
OS Plasmodium falciparum (Isolate 3D7).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=36329;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=3D7;  
RX MEDLINE=21036610; PubMed=11163452;  
RA Spielmann T., Beck H.P.;  
RT "Analysis of stage-specific transcription in Plasmodium falciparum reveals a set of genes exclusively transcribed in ring stage parasites.";  
RT Mol. Biochem. Parasitol. 111:453-458(2000).  
RL EMBL; AJ290921; CAB92930.1; -  
DR Hypothetical protein.  
KW NON\_TER  
FT SEQUENCE 40 AA; 4758 MW; A6C6CAD52BAC1903 CRC64;

Query Match 100.0%; Score 21; DB 5; Length 40;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
Db 23 SYDA 26

RESULT 2  
Q8YX6 PRELIMINARY; PRT; 53 AA.  
ID Q8YX6;  
AC Q8YX6;  
DT 01-MAR-2002 (TREMBlrel. 20, Created)

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Hypothetical protein Asl0716.  
 GN Asl0716.  
 OS Anabaena sp. (strain PCC 7120).  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
 OX NCBI\_TaxID=103690;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21595285; PubMed=11759840;  
 RA Kaneke T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
 RA Watanabe A., Iriyuchi M., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kashiida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,  
 RA Yasuda M., Tabata S.;  
 RT "Complete genomic sequence of the filamentous nitrogen-fixing  
 RT cyanobacterium Anabaena sp. strain PCC 7120.";  
 RL DNA Res. 8:205-213(2001).  
 DR EMBL; AF003583; BAB72673.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 53 AA; 6029 MW; 42F9CC5A554CF40 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 53;  
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 DB 2 SYDA 5

RESULT 3  
 ID Q12081 PRELIMINARY; PRT; 56 AA.  
 AC Q12081;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Hypothetical protein (fragment).  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX PLASMID=2-MICRON;  
 RX MEDLINE=87120291; PubMed=3027892;  
 RA Parikh V.S., Morgan M.M., Scott R., Clements L.S., Butow R.A.;  
 RT "The mitochondrial genotype can influence nuclear gene expression in  
 RT yeast.";  
 RL Science 235:576-580(1987).  
 DR EMBL; M15168; AAA66931.1; -;  
 DR EMBL; M15167; AAA66930.1; -;  
 KW Hypothetical protein.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 56 AA; 6510 MW; 6F9BF49FB86FE60B CRC64;

Query Match 100.0%; Score 21; DB 3; Length 56;  
 Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 DB 52 SYDA 55

RESULT 4  
 ID Q9X374 PRELIMINARY; PRT; 67 AA.  
 AC Q9X374;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE PX01-105.  
 OS Bacillus anthracis.  
 OC Plasmid virulence plasmid PX01.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
 OC Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1392;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=STERNE;  
 RX MEDLINE=99445483; PubMed=10515943;  
 RA Okinaka R.T., Cloud K., Hampton O., Hoffmaster A.R., Hill K.K.,  
 RA Klein P., Koehler T.M., Lanke G., Kumano S., Mahillon J., Manter D.,  
 RA Martinez Y., Ricke D., Svensson R., Jackson P.D.;  
 RT "Sequence and organization of PX01, the large Bacillus anthracis  
 RT plasmid harboring the anthrax toxin genes.";  
 RL J. Bacteriol. 181:6509-6515(1999).  
 DR EMBL; AF065404; AAD32409.1; -;  
 KW Plasmid.  
 SQ SEQUENCE 67 AA; 7660 MW; FDCD7CF784148FCB CRC64;

Query Match 100.0%; Score 21; DB 2; Length 67;  
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 DB 25 SYDA 28

RESULT 5  
 ID Q9RIC5 PRELIMINARY; PRT; 75 AA.  
 AC Q9RIC5;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Hypothetical protein YPM1.68A.  
 GN YPM1.68A.  
 OS Yersinia pestis.  
 OC Plasmid pMT1 (pMT-1).  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Yersinia.  
 OX NCBI\_TaxID=632;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=CO-92 / BIOVAR ORIENTALIS;  
 RX MEDLINE=21470413; PubMed=11586360;  
 RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,  
 RA Prentice M.B., Sepahinia M., James K.D., Churcher C., Mungall K.L.,  
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,  
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,  
 RA Fellwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,  
 RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,  
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;  
 RT "Genome sequence of Yersinia pestis, the causative agent of plague.";  
 RL Nature 413:523-527(2001).  
 DR EMBL; AL117211; CAB62375.1; -;  
 KW Hypothetical protein; Plasmid; Complete proteome.  
 SQ SEQUENCE 75 AA; 8293 MW; C6C26AFC212F2A8F CRC64;

Query Match 100.0%; Score 21; DB 16; Length 75;  
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 DB 36 SYDA 39

RESULT 6  
 ID Q9MC41 PRELIMINARY; PRT; 77 AA.  
 AC Q9MC41



AC Q9MC41;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Orf91.  
 GN ORF91.  
 OS Bacteriophage D3.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;  
 OC Lambda-like viruses.  
 OC NCBI\_TaxID=31535;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20042341; PubMed=10572124;  
 RA Gilakjan Z.A., Kropinski A.M.;  
 RT "Cloning and analysis of the capsid morphogenesis genes of Pseudomonas  
 RT aeruginosa bacteriophage D3; another example of protein chain mail?";  
 RL J. Bacteriol. 181:7221-7227(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20485557; PubMed=11029426;  
 RA Kropinski A.M.;  
 RT "Sequence of the Genome of the Temperate, Serotype-Converting,  
 RT Pseudomonas aeruginosa Bacteriophage D3";  
 RL J. Bacteriol. 182:6066-6074(2000).  
 DR EMBL; AF165214; AAF80767.1;  
 SQ SEQUENCE 77 AA; 8266 MW; 39DAE91077F7234B CRC64;

Query Match 100.0%; Score 21; DB 9; Length 77;  
 Best Local Similarity 100.0%; Pred. No. 5.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 DB 64 SYDA 67

RESULT 7  
 09Z3S8 PRELIMINARY; PRT; 78 AA.  
 AC 09Z3S8;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Ru1B (Fragment).  
 GN Ru1B  
 OS Pseudomonas syringae (pv. tomato).  
 OC Plasmid pPT23A.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OC NCBI\_TaxID=323;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PT23;  
 RX MEDLINE=99173737; PubMed=10075415;  
 RA Gibbon M.J., Sesma A., Canal A., Wood J.R., Hidalgo E., Brown J.,  
 RA Vivian A., Murrillo J.;  
 RT "Replication regions from plant-pathogenic Pseudomonas syringae  
 RT plasmids are similar to ColE2-replicon.";  
 RL Microbiology 145:325-334(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PT23;  
 RX MEDLINE=97080503; PubMed=8921848;  
 RA Sundin G.W., Kidambi S.P., Ullrich M., Bender C.L.;  
 RT "Resistance to ultraviolet light in Pseudomonas syringae: sequence and  
 RT functional analysis of the plasmid-encoded ru1AB genes.";  
 RL Gene 177:77-81(1996).  
 DR EMBL; AJ245410; CAB37318.1;  
 DR InterPro; IPR001126; UMOG\_1like.  
 DR Pfam; PF00817; IMS; 1.  
 KW Plasmid.  
 FT NON\_TER 78 78  
 SQ SEQUENCE 78 AA; 8531 MW; DAF650CB225065A7 CRC64;

Query Match 100.0%; Score 21; DB 2; Length 78;  
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 DB 53 SYDA 56

RESULT 8  
 09Z7E0 PRELIMINARY; PRT; 80 AA.  
 AC 09Z7E0;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Hypothetical protein lin2849.  
 GN LIN2849.  
 OS Listeria innocua.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
 OC Listeriaceae; Listeria.  
 OC NCBI\_TaxID=1642;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CLIP 11262 / SEROVAR 6A;  
 RX PubMed=11679669;  
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,  
 RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,  
 RA Charlot A., Chetoui F., Couve E., de Daruvar A., Dehoux P.,  
 RA Domann K.E., Dominguez-Bernal G., Duchaud E., Durant L., Dussauget O.,  
 RA Ertian K.-D., Fajh H., Garcia-del Portillo F., Garrido P.,  
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,  
 RA Jones L.-M., Kaerst U., Kreft J., Kunz M., Kunst F., Kurapkut G.,  
 RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,  
 RA Nordstok G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,  
 RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,  
 RA Vazquez-Boland J.-A., Voss H., Wehlund J., Cossart P.;  
 RT "Comparative genomics of Listeria species.";  
 RL Science 294:849-852(2001).  
 DR EMBL; AL596173; CAC98075.1;  
 DR L1stlist; LIN02849;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 80 AA; 9908 MW; B625A9DF7CF7CAD6 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 80;  
 Best Local Similarity 100.0%; Pred. No. 5.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 DB 22 SYDA 25

RESULT 9  
 09ZGZ2 PRELIMINARY; PRT; 81 AA.  
 AC 09ZGZ2;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
 DE Hypothetical 9.0 kDa protein.  
 GN Y1075.  
 OS Yersinia pestis.  
 OC Plasmid pMT-1.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Yersinia.  
 OC NCBI\_TaxID=632;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=KIM10+;  
 RX MEDLINE=99043898; PubMed=9826348;  
 RLindler L.E., Plano G.V., Burland V., Mayhew G.F., Blattner F.R.;

RT "Complete DNA sequence and detailed analysis of the *Yersinia pestis*  
 RT KIM5 plasmid encoding murine toxin and capsular antigen.",  
 RL Infect. Immun. 66:5731-5742(1998).  
 DR EMBL: AF074611; AAC82735.1;  
 KM Hypothetical protein; Plasmid.  
 SO SEQUENCE 81 AA; 8968 MW; 4369EB64DAF9C093 CRC64;

Query Match 100.0%; Score 21; DB 2; Length 81;  
 Best Local Similarity 100.0%; Pred. No. 5.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SYDA 4  
 Db 42 SYDA 45

RESULT 10  
 Q9K510 PRELIMINARY; PRT; 81 AA.  
 AC Q9K510;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Hypothetical 9.3 kDa protein.  
 OS *Pseudomonas syringae* (pv. *pisi*).  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC *Pseudomonas*.  
 OX NCBI\_TaxID=59510;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=RACE 2;  
 RA Arnold D.L., Jackson R.W., Vivian A.;  
 RT "Evidence for the mobility of an avirulence gene, *avrPpiA*, between  
 RT the chromosome and plasmids of races of *Pseudomonas syringae* pv.  
 RT *pisi*.";  
 RL Mol. Plant Pathol. 1:195-199(2000).  
 DR EMBL: AJ251482; CAB96971.1; -  
 DR InterPro: IPR001126; UMC1-like.  
 DR Pfam: PF00817; IMS; 1.  
 KW Hypothetical protein.  
 SO SEQUENCE 81 AA; 9277 MW; 0EFD90A9398DDF3F CRC64;

Query Match 100.0%; Score 21; DB 2; Length 81;  
 Best Local Similarity 100.0%; Pred. No. 5.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SYDA 4  
 Db 47 SYDA 50

RESULT 11  
 Q8Y3X8 PRELIMINARY; PRT; 81 AA.  
 AC Q8Y3X8;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Hypothetical protein lmo2701.  
 GN lmo2701.  
 OS *Listeria monocytogenes*.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
 OC Listeriaceae; Listeria.  
 OX NCBI\_TaxID=1639;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=EGD-E / SEROVAR 1/2A;  
 RX MEDLINE=21537279; PubMed=11679669;  
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,  
 RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,  
 RA Charbit A., Chetoui F., Couve E., de Daruvar A., Denoux P.,  
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussauget O.,  
 RA Ertlan K.-D., Feihl H., Garcia-del Portillo F., Garrido P.,

RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,  
 RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst J., Kurupkat G.,  
 RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjati H.,  
 RA Nordstedt G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,  
 RA Remmel B., Rose M., Schlutener T., Simoes N., Tierrez A.,  
 RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;  
 RT "Comparative genomics of *Listeria species*."  
 RL Science 294:849-852(2001).  
 DR EMBL: AL591984; CAD00914.1; -  
 DR L1stlist; lmo02701; -  
 KM Hypothetical protein; Complete proteome.  
 SO SEQUENCE 81 AA; 10039 MW; C94155A9DB68A9B CRC64;

Query Match 100.0%; Score 21; DB 16; Length 81;  
 Best Local Similarity 100.0%; Pred. No. 5.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SYDA 4  
 Db 22 SYDA 25

RESULT 12  
 Q9V0T7 PRELIMINARY; PRT; 82 AA.  
 AC Q9V0T7;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE CG16713 protein.  
 GN CG16713.  
 OS *Drosophila melanogaster* (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amaratilake P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Fletter B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu I., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durrin K.J., Evangelista C.C., Ferrara C., Ferriera S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Gary N.S., Galardi W.M., Glaser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heilmann T.J., Hernandez J.R., Houck C.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalusz F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Liao X., Lel Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paoleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J.R., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Sytkas R., Tector C., Turner R., Ventle E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003579; AAF51077.1; -;  
 DR HSP: P02760; 1BK.  
 DR FLYBASE: FBgn0031560; CG16713.  
 DR InterPro: IPR002223; Kunitz\_BPTI.  
 DR Pfam: PF00014; Kunitz\_BPTI; 1.  
 DR PRINTS: PR00759; BASICPTASE.  
 DR ProDom: PD000222; Kunitz\_BPTI; 1.  
 DR SMART: SM00131; KU; 1.  
 DR PROSITE: PS00280; BPTI\_KUNITZ\_1; 1.  
 DR PROSITE: PS00279; BPTI\_KUNITZ\_2; 1.  
 KW Serine protease inhibitor.  
 SQ SEQUENCE 82 AA; 9078 MW; 758DD1F9A830022B CRC64;

Query Match 100.0%; Score 21; DB 5; Length 82;  
 Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 DB 47 SYDA 50

## RESULT 13

082021 PRELIMINARY; PRT; 84 AA.

AC 082021;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DE Putative cytoplasmic protein.  
 GN YCEP OR STM1161.  
 OS *Salmonella typhimurium*.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC *Salmonella*.  
 OX NCBI\_TaxID=602;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-LT2 / SGC1412 / ATCC 700720;  
 RX MEDLINE=21534948; PubMed=11677609;  
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,  
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
 RA Waterston R., Wilson R.K.;  
 RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium  
 LT2.";  
 RL Nature 413:852-856(2001).  
 DR EMBL: AE008750; AAL20091.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 84 AA; 9717 MW; 078AF3C835A342F CRC64;

Query Match 100.0%; Score 21; DB 16; Length 84;  
 Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 DB 23 SYDA 26

## RESULT 14

08X8N9 PRELIMINARY; PRT; 84 AA.

AC 08X8N9;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Orf, hypothetical protein.

GN YCEP OR 21697 OR ECS1438.  
 OS *Escherichia coli* O157:H7.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC *Escherichia*.  
 OX NCBI\_TaxID=83334;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
 RX MEDLINE=21074935; PubMed=11206551;

RA Berna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Postal G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potomousis K.,  
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";  
 RL Nature 409:529-533(2001).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / RMD 0509952;  
 RX MEDLINE=21156231; PubMed=11258796;  
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,  
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;  
 RT "Complete genome sequence of enterohaemorrhagic *Escherichia coli*  
 O157:H7 and genomic comparison with a laboratory strain K-12.";  
 RL DNA Res. 8:11-22(2001).  
 DR EMBL: AE005316; AAG55806.1; -;  
 DR EMBL: AP002555; BAB34861.1; -;  
 KW Complete proteome.  
 SQ SEQUENCE 84 AA; 9633 MW; EB7103C964E5D5FE CRC64;

Query Match 100.0%; Score 21; DB 16; Length 84;  
 Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 DB 23 SYDA 26

## RESULT 15

068916 PRELIMINARY; PRT; 85 AA.

AC 068916;  
 DT 01-AUG-1998 (TREMBLrel. 07, Created)  
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Acyl carrier protein.  
 GN FRN1.  
 OS *Streptomyces roseofulvus*.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=33902;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=94237486; PubMed=8181754;  
 RA Bibb M.J., Sherman D.H., Omura S., Hopwood D.A.;  
 RT "Cloning, sequencing and deduced functions of a cluster of  
 RT Streptomyces genes probably encoding biosynthesis of the polyketide  
 RT antibiotic frenolicin.";  
 RL Gene 142:31-39(1994).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RA Reeves C.D., Soliday C.L.;  
 RT "Analysis of a 27 kb region of *Streptomyces roseofulvus* containing  
 RT genes for frenolicin biosynthesis.";  
 RL Submitted (APR-1998) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF058302; AAC18105.1; -;  
 DR InterPro: IPR003880; Pantine\_attach.  
 DR Pfam: PF00550; pp-binding; 1.  
 DR PROSITE: PS50075; ACP\_DOMAIN; 1.

KW Phosphopantetheine.  
 SQ SEQUENCE 85 AA; 9412 MW; B2615248D312862A CRC64;  
 Query Match 100.0%; Score 21; DB 2; Length 85;  
 Best Local Similarity 100.0%; Pred. No. 6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 Db 16 SYDA 19

RESULT 16  
 Q827L4 PRELIMINARY; PRT; 85 AA.  
 ID Q827L4;  
 AC Q827L4;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Hypothetical protein STY1199.  
 GN STY1199.  
 OS Salmomella typhi.  
 CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 CC Salmomella.  
 OX NCBI\_TaxID=601;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=21534947; PubMed=11677608;  
 RX Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J., Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M., Baker S., Basham D., Brooks K., Chillingworth T., Connerton P., Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J., Feltham T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jagels K., Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C., Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;  
 RT "Complete genome sequence of a multiple drug resistant Salmomella enterica serovar Typhi CT18.";  
 RL Nature 413:848-852(2001).  
 DR EMBL: AL627269; CAD08285.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 85 AA; 9829 MW; BAD68EC84B6CB561 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 85;  
 Best Local Similarity 100.0%; Pred. No. 6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 Db 24 SYDA 27

RESULT 17  
 Q9GY45 PRELIMINARY; PRT; 87 AA.  
 ID Q9GY45;  
 AC Q9GY45;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Hypothetical predicted protein LM12.105, unknown function.  
 GN LM12.105.  
 OS Leishmania major.  
 CC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
 OX NCBI\_TaxID=5664;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-FRIEDLIN;  
 RA Murphy L., Quail M., Harris D., Rajandream M., Ivens A., Barrell B., Oliver K.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL390114; CAC02005.2; -;  
 DR InterPro; IPR003089; AB\_hydrolase.

DR InterPro; IPR000379; Ser\_estrs\_site.  
 DR PRINTS; PR00111; ABHYDROLASE.  
 SQ SEQUENCE 87 AA; 9449 MW; 7F983BD405F9E121A CRC64;  
 Query Match 100.0%; Score 21; DB 5; Length 87;  
 Best Local Similarity 100.0%; Pred. No. 6.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 Db 22 SYDA 25

RESULT 18  
 Q72135 PRELIMINARY; PRT; 92 AA.  
 ID Q72135;  
 AC Q72135;  
 DT 01-AUG-1998 (TREMBLrel. 07, Created)  
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE L1 protein (Fragment).  
 OS Human papillomavirus.  
 CC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 CC Papillomavirus.  
 OX NCBI\_TaxID=10566;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-HPVX24;  
 RX MEDLINE=20295075; PubMed=10834958;  
 RA Berkhout R.J.M., Bowes Bavinck J.N., ter Schegget J., RT "Persistence of Human Papillomavirus DNA in Benign and (Pre)malignant Skin Lesions from Renal Transplant Recipients.";  
 RL J. Clin. Microbiol. 38:2087-2096(2000).  
 DR EMBL: AF054881; AAC12747.1; -;  
 DR InterPro; IPR002210; PV\_capsid\_L1.  
 DR Pfam; PF005007; late\_protein\_L1;  
 DR ProDom; PD000544; PV\_capsid\_L1; 1.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 92 AA; 10620 MW; BFBAFF632FA59D07 CRC64;

Query Match 100.0%; Score 21; DB 12; Length 92;  
 Best Local Similarity 100.0%; Pred. No. 6.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 Db 15 SYDA 18

RESULT 19  
 Q91241 PRELIMINARY; PRT; 93 AA.  
 ID Q91241;  
 AC Q91241;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Hypothetical 10.5 kDa protein (Fragment).  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC010220; AAH10220.1; -;  
 DR InterPro; IPR002005; Rab\_GDI\_REP.  
 DR Pfam; PF00996; GDI; 1.  
 KW Hypothetical protein.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 93 AA; 10537 MW; 74DBAF69BBAB5FA6 CRC64;

Query Match 100.0%; Score 21; DB 11; Length 93;  
 Best Local Similarity 100.0%; Pred. No. 6.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 DB 49 SYDA 52

## RESULT 20

ID 085877 PRELIMINARY; PRT; 95 AA.  
 AC 085877;  
 DT 01-NOV-1998 (TReMBLrel. 08, Created)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
 DE Hypothetical 10.5 kDa protein.  
 GN ORF392.  
 OS Sphingomonas aromaticivorans.  
 OC Plasmid pNL1.  
 OC Bacteria; Proteobacteria; alpha subdivision; Sphingomonadaceae;  
 OC Novosphingobium.  
 OX NCBI\_TaxID=48935;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-F199;  
 RA Romine M.F., Stillwell L.C., Wong K.-K., Thurston S.J., Sisk E.C.,  
 RA Sensen C.W., Gaasterland T., Saffier J.D., Fredrickson J.K.;  
 RT "Complete sequence of a 184 kb catabolic plasmid from Sphingomonas  
 RT aromaticivorans strain F199."  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF079317; AAD03892.1; -  
 KW Hypothetical protein; Plasmid.  
 SQ SEQUENCE 95 AA; 10535 MW; AC530A9B1563846D CRC64;

Query Match 100.0%; Score 21; DB 2; Length 95;  
 Best Local Similarity 100.0%; Pred. No. 6.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 DB 57 SYDA 60

## RESULT 21

ID 08Y0B6 PRELIMINARY; PRT; 96 AA.  
 AC 08Y0B6;  
 DT 01-MAR-2002 (TReMBLrel. 20, Created)  
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
 DE Hypothetical protein RSC1128.  
 GN RSC1128 OR RS04625.  
 OS Ralstonia solanacearum (Pseudomonas solanacearum).  
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;  
 OC Ralstonia.  
 OX NCBI\_TaxID=305;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-GM11000;  
 RC MEDLINE-21681879; PubMed-11823852;  
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,  
 RA Ariat M., Billault A., Brottier P., Camus J.C., Catolico L.,  
 RA Chandler M., Choisme N., Clandel-Renard C., Cunac S., Demange N.,  
 RA Gaspier C., Laveit M., Moisan A., Robert C., Saurin W., Schlex T.,  
 RA Sigulter P., Thebaud P., Whalen M., Wincker P., Levy M.,  
 RA Weisenbach J., Boucher C.A.;  
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum."  
 RT Nature 415:497-502(2002).  
 RL EMBL: AL646062; CADI4830.1; -  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 96 AA; 10601 MW; 960D8A89DB418D60 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 96;  
 Best Local Similarity 100.0%; Pred. No. 6.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 DB 62 SYDA 65

## RESULT 22

ID 014371 PRELIMINARY; PRT; 99 AA.  
 AC 014371;  
 DT 01-JAN-1998 (TReMBLrel. 05, Created)  
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
 DE 60S ribosomal protein L7B (Fragment).  
 OS Schizosaccharomyces pombe (Fission Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972H-;  
 RA Jang Y.-J., Yoo H.-S.;  
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U97365; AAB63857.1; -  
 DR InterPro: IPR000517; Ribosomal\_L30.  
 DR Pfam: PF00327; Ribosomal\_L30; 1.  
 KW Ribosomal protein.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 99 AA; 11302 MW; 01F060749ED147BD CRC64;

Query Match 100.0%; Score 21; DB 3; Length 99;  
 Best Local Similarity 100.0%; Pred. No. 7.1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
 ||||  
 DB 52 SYDA 55

## RESULT 23

ID 08S0G9 PRELIMINARY; PRT; 100 AA.  
 AC 08S0G9;  
 DT 01-JUN-2002 (TReMBLrel. 21, Created)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE B1142C05.31 protein.  
 GN B1142C05.31.  
 OS Oryza sativa (Japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Eriocaridaceae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa (Japonica cultivar-group) genomic DNA, chromosome 1, BAC  
 RT clone:B1142C05."  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AP003410; BAB89131.1; -  
 SQ SEQUENCE 100 AA; 11168 MW; 0E721D6E0C06CA CRC64;

Query Match 100.0%; Score 21; DB 10; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 7.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4

Db 75 SYDA 78  
||||

## RESULT 24

Q8RH72 PRELIMINARY; PRT; 100 AA.  
ID Q8RH72;  
AC Q8RH72;  
DT 01-JUN-2002 (TREMblrel. 21, Created)  
DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
DE Hypothetical protein FN0038.  
GN FN0038.  
OS Fusobacterium nucleatum (subsp. nucleatum).  
OC Bacteria; Fusobacteria; Fusobacterium.  
OX NCBI\_TaxID=76856;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 25586;  
RX MEDLINE=21886394; PubMed=11889109;  
RA Kapralat V., Anderson I., Ivanova N., Reznik G., Los T., Lykdis A.,  
RA Bhattacharyya A., Barton A., Gardner W., Grecklin G., Zhu L.,  
RA Vasiera O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,  
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,  
RA Farnstein M., Kyrides N., Overbeek R.;  
RT "Genome sequence and analysis of the oral bacterium Fusobacterium  
RT nucleatum strain ATCC 25586.";  
RL J. Bacteriol. 184:2005-2018(2002).  
DR EMBL: AE010519; AAL94251.1;-  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 100 AA; 11550 MW; 0A65A09071054742 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 100;  
Best Local Similarity 100.0%; Pred. No. 7.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SYDA 4  
||||  
Db 63 SYDA 66

## RESULT 25

085916 PRELIMINARY; PRT; 103 AA.  
ID 085916  
AC 085916;  
DT 01-NOV-1998 (TREMblrel. 08, Created)  
DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)  
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
DE Hypothetical 11.4 kDa protein.  
GN ORF680.  
OS Sphingomonas aromaticaivorans.  
OG Plasmid pNL1.  
OC Bacteria; Proteobacteria; alpha subdivision; Sphingomonadaceae;  
OC Novosphingobium.  
OX NCBI\_TaxID=48935;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-F199;  
RA Romine M.E., Stillwell L.C., Wong K.-K., Thurston S.J., Sisk E.C.,  
RA Samsen C.W., Gaasterland T., Saffier J.D., Fredrickson J.K.;  
RT "Complete sequence of a 184 kb catallolic plasmid from Sphingomonas  
RT aromaticaivorans strain F199.";  
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF079317; AAD03931.1;-  
KW Hypothetical protein; Plasmid.  
SQ SEQUENCE 103 AA; 11444 MW; 489F760C930BF103 CRC64;

Query Match 100.0%; Score 21; DB 2; Length 103;  
Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SYDA 4  
||||

Db 17 SYDA 20  
||||

## RESULT 26

09A108 PRELIMINARY; PRT; 106 AA.  
ID 09A108;  
AC 09A108;  
DT 01-JUN-2001 (TREMblrel. 17, Created)  
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)  
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
DE Hypothetical protein SPY0147.  
GN SPY0147.  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
OC Streptococcaceae; Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SF370 / ATCC 700294 / SEROTYPE M1;  
RX MEDLINE=21192684; PubMed=11296296;  
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,  
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,  
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,  
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;  
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).  
DR EMBL: AE006484; AAK33252.1;-  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 106 AA; 12466 MW; 0743A62E08B19951 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 106;  
Best Local Similarity 100.0%; Pred. No. 7.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SYDA 4  
||||  
Db 21 SYDA 24

## RESULT 27

09HN89 PRELIMINARY; PRT; 106 AA.  
ID 09HN89  
AC 09HN89;  
DT 01-MAR-2001 (TREMblrel. 16, Created)  
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)  
DE Vng2202n.  
GN VNG2202H.  
OS Halobacterium sp. (strain NRC-1).  
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;  
OC Halobacteriaceae; Halobacterium.  
OX NCBI\_TaxID=64091;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20504483; PubMed=11016950;  
RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,  
RA Shukla H.D., Lasky S.R., Ballig N.S., Thorsson V., Sbrogna J.,  
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,  
RA Leitbauer B., Keller K., Cruz R., Danson M.O., Hough D.W.,  
RA Madocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,  
RA Isenbarger T.A., Peck R.F., Pohlischroder M., Spudis J.L., Jung K.-H.,  
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;  
RT "Genome sequence of Halobacterium species NRC-1.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
DR EMBL: AE005107; AAG20332.1;-  
KW Complete proteome.  
SQ SEQUENCE 106 AA; 10605 MW; 66B4FB15BF9259 CRC64;

Query Match 100.0%; Score 21; DB 17; Length 106;  
Best Local Similarity 100.0%; Pred. No. 7.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
|||||  
DB 78 SYDA 81

## RESULT 28

ID O8W5F8 PRELIMINARY; PRT: 109 AA.

AC O8W5F8;  
DT 01-MAR-2002 (TREMBlrel. 20, Created)  
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Hypothetical 11.2 kDa protein (Putative rapid alkalization factor protein).  
GN OSJNB0008A05.5 OR OSJNB0073120.23.

OS Oryza sativa (Rice), and  
OS Oryza sativa (Japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Eriarthroideae; Oryzeae; Oryza.  
OX NCBI\_TaxID=4530, 39947;

RP SEQUENCE FROM N.A.

RC STRAIN=CV. NIPPONBARE;  
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Kim M.M.,  
RA Overton II L.L., Bera J.J., Tsirlin T., Krol M.I., Jatrani B.B.,  
RA Jin S.S., Koo H., Zismann V., Hsiao J., Blunt S., Vanaken S.S.,  
RA Uteback T.T., Feldblyum T.V., Yang Q.Q., Haas B.J., Suh B.B.,  
RA Peterson J.J., Quackenbush J., White O., Salzberg S.L., Fraser C.M.,  
RT "Oryza sativa chromosome 10 BAC OSJNB0008A05 genomic sequence."  
RT Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.

RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.

RC STRAIN=CV. NIPPONBARE;  
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Kim M.M.,  
RA Overton II L.L., Bera J.J., Tsirlin T., Krol M.I., Jatrani B.B.,  
RA Jin S.S., Koo H., Zismann V., Hsiao J., Blunt S., Vanaken S.S.,  
RA Uteback T.T., Feldblyum T.V., Yang Q.Q., Haas B.J., Suh B.B.,  
RA Peterson J.J., Quackenbush J., White O., Salzberg S.L., Fraser C.M.,  
RT "Oryza sativa chromosome 10 BAC OSJNB0073120 genomic sequence."  
RT Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AC091749; ALU31098.1;  
DR EMBL: AC091774; AAM19046.1;  
KW Hypothetical protein.  
SQ SEQUENCE 109 AA; 11235 MW; EA4CC9DC03D812A7 CRC64;

Query Match 100.0%; Score 21; DB 10; Length 109;

Best Local Similarity 100.0%; Pred. No. 7.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
|||||  
DB 67 SYDA 70

## RESULT 29

ID O942U5 PRELIMINARY; PRT: 111 AA.

AC O942U5;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE P0506E04.27 protein.  
GN P0506E04.27.

OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Eriarthroideae; Oryzeae; Oryza.  
OX NCBI\_TaxID=4530;

RP SEQUENCE FROM N.A.  
RC STRAIN=CV. NIPPONBARE;  
RA Sasaki T., Matsumoto T., Yamamoto K.;

RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC  
RT clone:P0506E04."  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AP003272; BAB67949.1;  
DR InterPro: IPR002086; Aldehyde\_dehydr.  
DR PROSITE: PS00070; ALDEHYDE\_DEHYDR\_CYS; UNKNOWN.1.  
SQ SEQUENCE 111 AA; 11334 MW; 74AB4291D5334A2A CRC64;

Query Match 100.0%; Score 21; DB 10; Length 111;

Best Local Similarity 100.0%; Pred. No. 8e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
|||||  
DB 69 SYDA 72

## RESULT 30

ID O9NBJ7 PRELIMINARY; PRT: 112 AA.

AC O9NBJ7;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Ribosomal protein L1 (Fragment).  
GN RPL-10A.

OS Caenorhabditis remanei.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditidae;  
OC Rhabdilitidae; Pelodierinae; Caenorhabditis.

OX NCBI\_TaxID=31234;

RP SEQUENCE FROM N.A.

RX MEDLINE=20428552; PubMed=10970881;  
RA Mitrovich Q.M., Anderson P.;  
RT "Unproductively spliced ribosomal protein mRNAs are natural targets of  
RT RNA surveillance in C. elegans."  
RL Genes Dev. 14:2173-2184(2000).  
CC -!- SIMILARITY: BELONGS TO THE L1P FAMILY OF RIBOSOMAL PROTEINS.

EMBL: AF247854; AAF7035.1;  
DR InterPro: IPR002143; Ribosomal\_L1.  
DR Pfam: PF00687; Ribosomal\_L1; 1.

DR PROSITE: PS01199; RIBOSOMAL\_L1; 1.  
KW Ribosomal protein.

FT NON\_TER 1 1  
FT NON\_TER 112 112

SQ SEQUENCE 112 AA; 12307 MW; F5CA6E1FECE6E01C CRC64;

Query Match 100.0%; Score 21; DB 5; Length 112;

Best Local Similarity 100.0%; Pred. No. 8.1e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYDA 4  
|||||  
DB 57 SYDA 60

## RESULT 31

ID O9XCB3 PRELIMINARY; PRT: 113 AA.

AC O9XCB3;  
DT 01-NOV-1999 (TREMBlrel. 12, Created)  
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Putative O-methyltransferase (Fragment).  
OS Rhodothermus marinus (Rhodothermus obanensis).

OC Bacteria; CF3 group; Rhodothermus group; Rhodothermus.  
OX NCBI\_TaxID=29549;

RP SEQUENCE FROM N.A.

RC STRAIN=ITI 376;  
RA Thorbjarnardottir E.T.T., Backman V.M., Blondal T.,  
RA Thorbjarnardottir S.H., Palsdottir A., Hauksdottir H.,  
RA Kristjansdottir S., Eggertsson G.;

```

RT "Heat shock in Rhodothermus marinus: Cloning and sequence analysis of
   the groEL, dnaK and dnaJ genes."
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF145250; AAD37972.1; -.
DR InterPro: IPR002935; Methyltransf_3.
DR Pfam: PF01596; Methyltransf_3; 1.
KW Methyltransferase; Transferase.
FT NON_TER
SQ SEQUENCE 113 AA; 12453 MW; D04A0E1D7DE15FF4 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 2; Length 113;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
   ||||
Db 41 SYDA 44

RESULT 32
Q95WK3 PRELIMINARY; PRT; 113 AA.
AC Q95WK3;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Glucose-6-phosphate dehydrogenase (Fragment).
GN G6PD.
OS Anopheles pseudopunctipennis.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Preyryota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
OC Anopheles.
OC NCBI_TaxID=46955;
RN [1]
RP SEQUENCE FROM N.A.
RA Krzyzanski J., Milerson R.C., Besansky N.J.;
RT "Toward understanding Anopheles (Diptera, Culicidae) phylogeny:
   insights from nuclear single copy genes and the weight of evidence.";
RL Syst. Biol. 50:540-556(2001).
DR EMBL: AF317810; AAL18429.1; -.
DR InterPro: IPR001282; G6PD.
DR Pfam: PF00479; G6PD; 1.
DR Prodom: PD001129; G6PD; 1.
FT NON_TER
FT 1
SQ SEQUENCE 113 AA; 13423 MW; DD117794D5784E77 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 5; Length 113;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
   ||||
Db 60 SYDA 63

RESULT 33
Q9RZP8 PRELIMINARY; PRT; 113 AA.
AC Q9RZP8;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Salicylate monooxygenase-related protein.
GN DRB0072.
OS Deinococcus radiodurans.
OC Plasmid MPI.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RI;

```

```

RX MEDLINE-20036896; PubMed-10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.D., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarewicz K.S., Aravind L., Daly M.J., Milton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
   radiodurans R1."
RL Science 286:1571-1577(1999).
DR EMBL: AE001826; AAF12594.1; -.
DR TIGR: DRB0072; -.
DR InterPro: IPR000733; Flav_monooxygenase.
DR InterPro: IPR003042; Ring_mnoxygenase.
DR Pfam: PF01360; Monooxygenase; 1.
DR PRINTS: PR00420; RINGMONOXGNASE.
KW Monooxygenase; Plasmid; Complete proteome.
SQ SEQUENCE 113 AA; 12561 MW; 9DE2C58D930A6D73 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 16; Length 113;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
   ||||
Db 104 SYDA 107

RESULT 34
Q8RSN6 PRELIMINARY; PRT; 115 AA.
AC Q8RSN6;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Upf54.8 protein.
GN UPF54.8.
OS uncultured bacterium.
OC Bacteria; environmental samples.
OC Bacteria; environmental samples.
OC NCBI_TaxID=77133;
RN [1]
RP SEQUENCE FROM N.A.
RA Tauch A.;
RT "The Incp beta plasmid pB4 encodes a tripartite antibiotic efflux
   system of the RND-MFP-OBP type conferring erythromycin and
   roxithromycin resistance in Pseudomonas sp. B13."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: A4431260; CAD24324.1; -.
KW Plasmid.
SQ SEQUENCE 115 AA; 12684 MW; D377032F611FB060 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 2; Length 115;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
   ||||
Db 93 SYDA 96

RESULT 35
Q92J33 PRELIMINARY; PRT; 115 AA.
AC Q92J33;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Hypothetical protein RC0236.
GN RC0236.
OS Rickettsia conorii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;

```



```

OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MALISH 7;
R MEDLINE=21442074; PubMed=11557893;
RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA Samsen D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA Reout D.;
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL Science 293:2093-2098(2001).
DR EMBL: AE008591; AL027774.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 115 AA; 13298 MW; C419CB796D5BD8A1 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 16; Length 115;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
   ||||
DB 87 SYDA 90

RESULT 36
Q991X7 PRELIMINARY; PRT; 116 AA.
AC Q991X7;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Hypothetical 12.8 kDa protein.
OS uncultured organism.
OC unclassified; environmental samples.
OX NCBI_TaxID=155900;
RN [1]
RP SEQUENCE FROM N.A.
RA Stokes H.W., Nield B.S., Mabbutt B.C., Nevalainen H., Holmes A.J.,
RA Gillings M.R.;
RT "Novel and diverse integrin-like gene cassettes are prevalent in
RT natural environments.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF349050; AK28563.1; -.
KW Hypothetical protein.
SQ SEQUENCE 116 AA; 12765 MW; D92EC31445B2BB6 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 14; Length 116;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
   ||||
DB 69 SYDA 72

RESULT 37
Q920S4 PRELIMINARY; PRT; 121 AA.
AC Q920S4;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative regulatory protein, arsr family.
GN RB1016 OR SMB21576.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Plasmid PSymb (megaplasmid 2).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396508; PubMed=11481431;

```

```

RA Finan T.M., Weidner S., Wong K., Buhmester J., Chain P.,
RA Vorhoeft F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouy J.,
RA Gidding B., Puehler A.;
RT "The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-
RT fixing endosymbiont Sinorhizobium meliloti.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
DR EMBL: AL603645; CAC49416.1; -.
DR InterPro: IPR001845; HTH_ArsR.
DR Pfam: PF01022; HTH_5; 1.
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 121 AA; 12870 MW; 466B316B40DE46C1 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 16; Length 121;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
   ||||
DB 76 SYDA 79

RESULT 38
Q9XVLO PRELIMINARY; PRT; 127 AA.
AC Q9XVLO;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE K09G1.2 protein.
GN K09G1.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Lloyd C.R.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=99069613; PubMed=9851916;
RX none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL: Z81101; CAB03198.1; -.
SQ SEQUENCE 127 AA; 14845 MW; F156F752A6C8E6B0 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 5; Length 127;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4
   ||||
DB 88 SYDA 91

RESULT 39
P73361 PRELIMINARY; PRT; 129 AA.
AC P73361;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical protein slr1210.
GN SLR1210.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,

```

RA Miyajima N., Hirose M., Sugita M., Sasamoto S., Kimura T.,  
 RA Hosouchi T., Matsuno A., Muraiki A., Nakazaki N., Naito K., Okumura S.,  
 RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,  
 RA Tabata S.,  
 RT "Sequence analysis of the genome of the unicellular cyanobacterium  
 RT *Synechocystis* sp. strain PCC6803. II. Sequence determination of the  
 RT entire genome and assignment of potential protein-coding regions.",  
 RL DNA Res. 3:109-136(1996).  
 DR EMBL: D90905; BAA17392.1; -  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 129 AA; 14936 MW; AD6EF19018322DEA CRC64;

Query Match 100.0%; Score 21; DB 16; Length 129;  
 Best Local Similarity 100.0%; Pred. No. 9.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SYDA 4  
 Db 28 SYDA 31

RESULT 40  
 ID 099WV3 PRELIMINARY; PRT; 130 AA.  
 AC 099WV3.  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Hypothetical protein SAV0273.  
 GN SAV0273 OR SA0262.  
 OS *Staphylococcus aureus* (strain Mu50 / ATCC 700699), and  
 OS *Staphylococcus aureus* (strain N315).  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
 OC *Staphylococcus*.  
 OX NCBI\_TaxID=158878, 158879;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-S. aureus (strain Mu50), and S. aureus (strain N315);  
 RX MEDLINE-21311952; PubMed-11418146;  
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,  
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Itoh T.,  
 RA Kanamori M., Matsunaru H., Maruyama A., Murakami H., Hosoyama A.,  
 RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,  
 RA Sekimizu K., Hiraoka H., Kuhara S., Goto S., Iabuzaki J.,  
 RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,  
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.,  
 RT "Whole genome sequencing of methicillin-resistant *Staphylococcus*  
 RT *aureus*."  
 RL Lancet 357:1225-1240(2001).  
 DR EMBL: AP003358; BAB56435.1; -  
 DR EMBL: AP003130; BAB41486.1; -  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 130 AA; 15765 MW; A3F692D13BD7B764 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 130;  
 Best Local Similarity 100.0%; Pred. No. 9.5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SYDA 4  
 Db 67 SYDA 70

RESULT 41  
 ID 09NBK3 PRELIMINARY; PRT; 132 AA.  
 AC 09NBK3.  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Ribosomal protein L1 (Fragment).  
 GN RPL-10A.  
 OS *Caenorhabditis briggsae*.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6238;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-20428552; PubMed-10970881;  
 RA Mitrovich Q.M., Anderson P.,  
 RT "Unproductively spliced ribosomal protein mRNAs are natural targets of  
 RT mRNA surveillance in *C. elegans*."  
 RL Genes Dev. 14:2173-2184(2000).  
 CC -I- SIMILARITY: BELONGS TO THE L1P FAMILY OF RIBOSOMAL PROTEINS.  
 DR EMBL: AF247848; AAF77029.1; -  
 DR InterPro: IPR002143; Ribosomal\_L1.  
 DR Pfam: PF00687; Ribosomal\_L1.  
 DR PROSITE: PS01199; RIBOSOMAL\_L1; 1.  
 KW Ribosomal protein.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 132 AA; 14838 MW; A7C168AE1895D807 CRC64;

Query Match 100.0%; Score 21; DB 5; Length 132;  
 Best Local Similarity 100.0%; Pred. No. 9.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SYDA 4  
 Db 77 SYDA 80

RESULT 42  
 ID 08X858 PRELIMINARY; PRT; 132 AA.  
 AC 08X858.  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Putative tail component of prophage CP-933K (Putative minor tail  
 DE protein).  
 GN Z0973 OR ECS0835.  
 OS *Escherichia coli* O157:H7.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC *Escherichia*.  
 OX NCBI\_TaxID=83334;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
 RX MEDLINE-21074935; PubMed-11206551;  
 RA Perna N.T., Plunkett G., III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Grodebeck E.J., Davis N.W., Lam A., Dimalanta E.T., Potamocis K.,  
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blattner F.R.,  
 RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7."  
 RL Nature 409:529-533(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / RIMD 0509952;  
 RX MEDLINE-21156231; PubMed-11258796;  
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,  
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.,  
 RT "Complete genome sequence of enterohaemorrhagic *Escherichia coli*  
 RT O157:H7 and genomic comparison with a laboratory strain K-12."  
 RL DNA Res. 8:11-22(2001).  
 DR EMBL: AF005257; AAG55129.1; -  
 DR EMBL: AP005253; BAB34258.1; ALT\_INIT.  
 KW Complete proteome.  
 SQ SEQUENCE 132 AA; 14744 MW; FB479F4F356B5C2 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 132;  
 Best Local Similarity 100.0%; Pred. No. 9.6e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
|||||

Db 96 SYDA 99

## RESULT 43

Q9XUK1 PRELIMINARY; PRT; 133 AA.

AC Q9XUK1;  
DT 01-NOV-1999 (TREMblrel. 12, Created)  
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)  
DE W07G1.4 protein.  
GN W07G1.4.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Percy C.M.;  
RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99069613; PubMed=9851916;  
RA none;  
RT "Genome sequence of the nematode C. elegans: A platform for  
investigating biology."  
RL Science 282:2012-2018(1998).  
DR EMBL: 282076; CAB04936.1; -.  
SQ SEQUENCE 133 AA; 15587 MW; 5D9F516D5C20BCB1 CRC64;

Query Match 100.0%; Score 21; DB 5; Length 133;  
Best Local Similarity 100.0%; Pred. No. 9.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
|||||

Db 87 SYDA 90

## RESULT 44

Q97HT8 PRELIMINARY; PRT; 133 AA.

AC Q97HT8;  
DT 01-OCT-2001 (TREMblrel. 18, Created)  
DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)  
DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
DE Phage related SSB-like protein.  
GN CAC1919.  
OS Clostridium acetobutylicum.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;  
OC Clostridiales; Clostridiaceae; Clostridium.  
OX NCBI\_TaxID=1488;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;  
RX MEDLINE=21359325; PubMed=11466286;  
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,  
Gibson R., Lee H.M., Dubois J., Qiu D., Hittl J., Wolf Y.I.,  
Bennett G.N., Koonin E.V., Smith D.R.;  
RA Tatunov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,  
RT "Genome sequence and comparative analysis of the solvent-producing  
bacterium Clostridium acetobutylicum."  
RL J. Bacteriol. 183:4823-4838(2001).  
DR EMBL: AE007698; AAK79882.1; -.  
DR InterPro: IPR000424; SSB\_protein.  
DR Pfam: PF00436; SSB.1.  
DR TIGRFAMs: TIGR00621; ssb.1.  
KW Complete proteome.  
SQ SEQUENCE 133 AA; 14980 MW; F0FC00CEBF938EDB CRC64;

Query Match 100.0%; Score 21; DB 16; Length 133;  
Best Local Similarity 100.0%; Pred. No. 9.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
|||||

Db 80 SYDA 83

## RESULT 45

Q97FH2 PRELIMINARY; PRT; 133 AA.

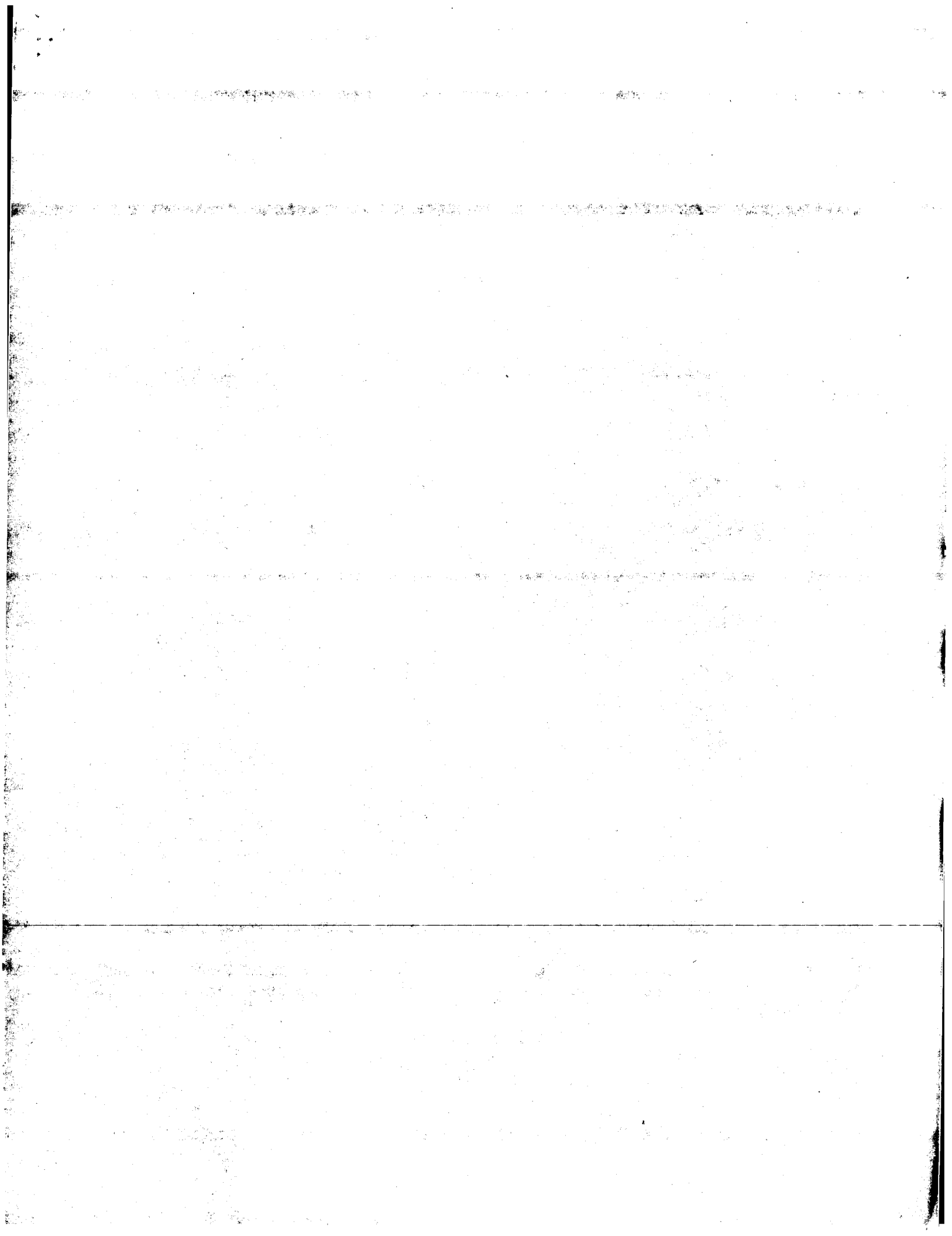
AC Q97FH2;  
DT 01-OCT-2001 (TREMblrel. 18, Created)  
DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)  
DE 01-OCT-2001 (TREMblrel. 18, Last annotation update)  
DE Uncharacterized conserved protein.  
GN CAC2767.  
OS Clostridium acetobutylicum.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;  
OC Clostridiales; Clostridiaceae; Clostridium.  
OX NCBI\_TaxID=1488;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;  
RX MEDLINE=21359325; PubMed=11466286;  
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,  
Gibson R., Lee H.M., Dubois J., Qiu D., Hittl J., Wolf Y.I.,  
RA Tatunov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,  
Bennett G.N., Koonin E.V., Smith D.R.;  
RT "Genome sequence and comparative analysis of the solvent-producing  
bacterium Clostridium acetobutylicum."  
RL J. Bacteriol. 183:4823-4838(2001).  
DR EMBL: AE007774; AAK80711.1; -.  
DR EMBL: AE007774; AAK80711.1; -.  
KW Complete proteome.  
SQ SEQUENCE 133 AA; 14776 MW; A607D42063381E1C CRC64;

Query Match 100.0%; Score 21; DB 16; Length 133;  
Best Local Similarity 100.0%; Pred. No. 9.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYDA 4  
|||||

Db 56 SYDA 59

Search completed: February 6, 2003, 11:21:58  
Job time : 24.3333 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 11:16:00 ; Search time 27.8333 Seconds  
(without alignments)  
19.150 Million cell updates/sec

Title: PAT943-4  
Perfect score: 18  
Sequence: 1 sida 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 2596

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A.Geneseq\_101002:\*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	18	100.0	7	16	AA72768	Mammalian ribonucle
2	18	100.0	7	19	AAW70154	Peptide produced b
3	18	100.0	7	19	AAW46010	Peptide #41 based
4	18	100.0	7	21	AAV96208	Arabidopsis AHAS s
5	18	100.0	7	21	AAV93850	Ribonucleotide red
6	18	100.0	7	22	AAE05612	N-terminal CT27 pe
7	18	100.0	8	18	AAW13438	Brain homing pepti
8	18	100.0	8	21	AAW07398	Brain homing pepti
9	18	100.0	8	21	AAV56002	Mucor hiemalis end
10	18	100.0	8	22	AAU68419	Human Breast cance

11	18	100.0	8	22	AAE11804	Phage peptide #12
12	18	100.0	8	23	ABG67883	Human ADP1 tryptic
13	18	100.0	8	23	AAU10715	Brain homing pepti
14	18	100.0	9	21	AAV93097	Transforming growt
15	18	100.0	9	22	AAV93639	Human hsp60 derive
16	18	100.0	10	16	AAW92570	Penicillium purpur
17	18	100.0	10	22	AAW61058	Peptide from Igy u
18	18	100.0	10	22	AAW68956	Saccharomyces cere
19	18	100.0	12	21	AAV92947	Transforming growt
20	18	100.0	12	21	AAV93007	Transforming growt
21	18	100.0	12	21	AAV93008	Transforming growt
22	18	100.0	12	21	AAV93009	Transforming growt
23	18	100.0	12	21	AAV93093	Transforming growt
24	18	100.0	13	19	AAW69624	Ste2 agonist from
25	18	100.0	13	20	AAW74421	Ste2 agonist M2 pe
26	18	100.0	13	21	AAW20756	Alpha-Mid-5 librat
27	18	100.0	13	22	AAW79149	Amino acid sequenc
28	18	100.0	13	22	AAW84497	Amino acid sequenc
29	18	100.0	14	21	AAV92950	Transforming growt
30	18	100.0	14	21	AAV93098	Transforming growt
31	18	100.0	14	22	AAW97105	Human peptide #380
32	18	100.0	14	22	AAW97480	Human peptide #735
33	18	100.0	15	16	AAW93618	HIV principal neut
34	18	100.0	16	21	AAV95420	Anti-angiogenic D3
35	18	100.0	19	14	AAW36604	N-terminus of 37 k
36	18	100.0	19	14	AAV91397	Human secreted pro
37	18	100.0	20	20	AAV36691	Fragment of human
38	18	100.0	21	14	AAW33356	Sequence of trypti
39	18	100.0	26	22	AAW51179	Human secreted pro
40	18	100.0	26	23	AAW86904	Insulin/insulin-11
41	18	100.0	29	18	AAW10854	Mab anti-HBSAg bin
42	18	100.0	30	16	AAW71050	Draculin peptide #
43	18	100.0	30	19	AAW31683	Clostridium perfr
44	18	100.0	32	21	AAV95418	Anti-angiogenic D3
45	18	100.0	33	18	AAW10869	Mab anti-HBSAg bin

#### ALIGNMENTS

RESULT 1  
AA72768  
ID AAR72768 standard; Protein; 7 AA.  
XX  
AC AAR72768;  
XX  
DF 02-FEB-1996 (first entry)  
XX  
DE Mammalian ribonucleotide reductase inhibitory peptide.  
XX  
KW Plasmodium falciparum; mammalian; ribonucleotide reductase; R2 subunit;  
KW C-terminal region; inhibitor; antimalarial therapy; Pf PR; parasite.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 1 /note="opt. acylated"  
XX  
XX  
PN WO9510526-A1.  
XX  
PD 20-APR-1995.  
XX  
PF 07-OCT-1994; 94WO-US11416.  
XX  
XX 14-OCT-1993; 93US-0136743.  
XX  
XX (UYPE-) UNIV PENNSYLVANIA.  
PA  
XX  
PI Cooperman BS, Fisher AL, Rubin H, Salem J;  
XX  
XX WPI; 1995-161730/21.  
XX

PT Plasmodium falciparum ribonucleotide reductase (Pf RR) sub-units -  
 PT also DNA and peptide(s) derived from sub-unit R2 C-terminus useful  
 PT in antimalarial compns as specific inhibitors of Pf RR activity  
 PS  
 XX Example 34 and 54; Page 61; 79pp; English.

CC The sequence of the peptide shown in AAR72739 corresponds to the  
 CC C-terminal sequence of both the mouse and the human ribonucleotide  
 CC reductase (RR) R2 subunit. The acylated form of the peptide is an  
 CC inhibitor of mammalian RR activity (IC50 = 10-20 microm on mouse  
 CC R2). Various peptides (including the present peptide, AAR72768) were  
 CC prepared in acylated and non-acylated form and assayed for inhibitory  
 CC effects on mammalian RR, compared to AAR72739. It was found that the  
 CC acylated form of the peptide given in AAR72740 was 8-16 times less  
 CC potent an inhibitor and was also found to specifically inhibit RR  
 CC activity of Plasmodium falciparum. It is thus a preferred peptide  
 CC of the invention, having antimalarial activity with little inhibitory  
 CC effect on mammalian (including human) ribonucleotide reductases.  
 CC The acylated form of the present peptide had an IC50 value of 40-42  
 CC microm against mouse recombinant RR R2.

CC  
 XX Sequence 7 AA;

Query Match 100.0%; Score 18; DB 16; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
 ||||  
 Db 2 SLDA 5

RESULT 2  
 AAM70154  
 ID AAM70154 standard; peptide; 7 AA.  
 XX  
 AC AAM70154;  
 XX  
 DT 07-DEC-1998 (first entry)  
 XX  
 DE Peptide produced by chymotrypsin/trypsin digestion of 45 kDa MPT 32.  
 XX  
 KM Glycopeptide: chymotrypsin; trypsin; reversed-phase HPLC; FAB-MS;  
 KM Fast Atom Bombardment-Mass Spectrometry; Edman degradation; mycobacteria;  
 KM prokaryote; antigen.  
 XX  
 OS Mycobacterium tuberculosis.  
 XX  
 PN WO9829132-A1.  
 XX  
 PD 09-JUL-1998.  
 XX  
 PF 29-DEC-1997; 97WO-US24189.  
 XX  
 PR 31-DEC-1996; 96US-0034003.  
 XX  
 PA (UYNV) UNTV NEW YORK STATE.  
 XX  
 PI Belisle JT, Laal S, Zolla-Pazner S;  
 DR WPI; 1998-387787/33.  
 XX  
 PT Early detection of mycobacterial infection - by testing a biological  
 PT fluid sample from a subject for the presence of antibodies reactive  
 PT with Mycobacterium tuberculosis antigens  
 XX  
 PS Example 3; Page 72; 170pp; English.

CC Sequences AAM70114, AAM70142-W70154 are N-terminal amino acid sequences  
 CC of peptides generated by chymotrypsin/trypsin digestion of the 45 kDa MPT  
 CC 32 glycoprotein. These sequences were obtained by purifying the 45 kDa  
 CC protein, digesting it with chymotrypsin/trypsin, and then separating the  
 CC products by reversed-phase HPLC on a column. All of the peptides were

CC subjected to both Fast Atom Bombardment-Mass Spectrometry (FAB-MS), and  
 CC automated Edman degradation to determine their molecular weight and  
 CC N-terminal amino acid sequence, respectively. The discovery of the  
 CC complete primary structure of MPT 32 glycoprotein enables further work  
 CC to be performed and the possibility of it being used as an antigen for  
 CC the early detection of mycobacterial disease.

CC  
 XX Sequence 7 AA;

Query Match 100.0%; Score 18; DB 19; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
 ||||  
 Db 4 SLDA 7

RESULT 3  
 AAM46010  
 ID AAM46010 standard; peptide; 7 AA.  
 XX  
 AC AAM46010;  
 XX  
 DT 03-JUL-1998 (first entry)  
 XX  
 DE Peptide #41 based on human SSTR 4 (residues 282-290).  
 XX  
 XX Hormone; receptor; antibody; vaccine; immunogen; somatostatin; IGF;  
 KM insulin-like growth factor binding protein; IGFBP; SSTR; diabetes;  
 KM somatostatin receptor; insulin-like growth factor.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 PN WO9744352-A1.  
 XX  
 PD 27-NOV-1997.  
 XX  
 PF 22-MAY-1997; 97WO-AU00312.  
 XX  
 PR 22-MAY-1996; 96AU-0009990.  
 XX  
 PA (NORT-) NORTHSTAR BIOLOGICALS PVT LTD.  
 XX  
 PI Gerraty NL, Kingston DJ, Westbrook SL;  
 XX WPI; 1998-018427/02.  
 XX  
 PT New non-naturally occurring peptide(s) - which are based on portions  
 PT of somatostatin, somatostatin receptors and insulin-like growth  
 PT factor binding protein  
 XX  
 PS Disclosure; Page 9; 136pp; English.  
 XX  
 CC Peptides AAM45983-W456025 are based on portions of somatostatin,  
 CC somatostatin receptors (SSTR) and insulin-like growth factor binding  
 CC proteins (IGFBP). They are capable of increasing weight gain, birth  
 CC weight, growth rates, milk production, levels of circulating insulin,  
 CC IGF-I and IGF-III, fibre production and muscle weight. They may be used  
 CC to modulate carbohydrate metabolism and in treatment of diabetes. The  
 CC oil carrier may be used for delivery of the peptides.

CC  
 XX Sequence 7 AA;

Query Match 100.0%; Score 18; DB 19; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
 ||||  
 Db 2 SLDA 5

## RESULT 4

AAV96208  
ID AAV96208 standard; Peptide; 7 AA.  
AC AAV96208;  
XX  
XX  
DT 11-AUG-2000 (first entry)  
DE Arabidopsis AHAS small subunit F2 gene fragment N-terminal sequence.  
XX  
XX Herbicide-resistance; acetohydroxy-acid synthase; AHAS;  
KW acetoacetate synthase; imidazolinone; sulfonyleurea;  
KM triazopyrimidine sulfonamide; sulfamoylurea; enzyme;  
KM pyrimidyl-oxy-benzoic acid; sulfonylecarboxamide; transgenic plant;  
KM branched-chain amino acid synthesis.  
XX  
XX Arabidopsis.  
OS  
PN W0200026390-A2.  
XX  
PD 11-MAY-2000.  
XX  
PF 28-OCT-1999; 99WO-US25452.  
XX  
PR 29-OCT-1998; 98US-0106239.  
XX  
PA (AMCY ) AMERICAN CYANAMID CO.  
XX  
PI Kakefuda G, Costello C, Sun M, Hu W;  
DR WPI; 2000-365633/31.  
XX  
PT New polynucleotide encoding eukaryotic acetohydroxy-acid synthetase  
PT small subunit protein for producing transgenic herbicide resistant  
PT plants and identifying mutations affecting enzymatic activity of the  
PT synthetase -  
XX  
XX  
PS Disclosure; Page 25; 57pp; English.

CC Arabidopsis acetohydroxy-acid synthase (AHAS) small subunit is needed  
CC for branched-chain amino acid synthesis and so is essential for life.  
CC AHAS is also known as acetoacetate synthase. Inhibition of this enzyme  
CC in plants would lead to plant death and therefore inhibitors would be  
CC potential herbicides. Certain herbicides are known to inhibit AHASs:  
CC imidazolinones, sulfonyleureas, triazopyrimidine sulfonamides,  
CC pyrimidyl-oxy-benzoic acids, sulfamoylureas and sulfonylecarboxamides.  
CC Mutant AHAS may be resistant to these herbicides and may be used to  
CC create herbicide resistant transgenic plants e.g. dicot and monocot crop  
CC plants. A thrombin cleavage site was incorporated into an AHAS small  
CC subunit gene. The AHAS small subunit gene fragments were cloned into  
CC plasmid expression vectors, and were expressed as glutathione  
CC transferase/AHAS small subunit fusion proteins. The glutathione  
CC transferase was then cleaved via the thrombin cleavage site. Cleavage  
CC however modifies the N-terminal sequence of the AHAS small subunit gene  
CC fragments, in that the glycine and serine residues of the cleavage site  
CC are maintained on the AHAS protein. The present sequence details the  
CC N-terminal sequence of one such AHAS small subunit gene fragment,  
CC F2, which is a near full length AHAS small subunit gene, with the  
CC modification due to cleavage.  
XX  
XX  
SQ Sequence 7 AA;

Query Match 100.0%; Score 18; DB 21; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|  
|  
|  
|  
DB 2 SLDA 5

## RESULT 5

AAV83850  
ID AAV83850 standard; peptide; 7 AA.  
XX  
XX  
AC AAV83850;  
XX  
XX  
DT 05-JUL-2000 (first entry)  
DE Ribonucleotide reductase inhibitor peptide #13.  
XX  
XX  
KW Cytostatic; antibacterial; virucide; antifungal; parasiticide; virus;  
KM antimetastatic; ribonucleotide reductase; peptide inhibitor; bacterium;  
KM DNA replication; tumour growth; metastasis; fungus; parasite.  
XX  
XX  
OS Synthetic.  
FH  
FT Key Location/Qualifiers  
FT Modified-site 1 /note="acetylated N-terminus"  
XX  
XX  
PN US6030942-A.  
XX  
PD 29-FEB-2000.  
XX  
PF 17-MAR-1998; 98US-0040216.  
XX  
PR 30-AUG-1996; 96US-0025146.  
PR 28-AUG-1997; 97US-0919748.  
XX  
XX  
PA (VYPE-) UNIV PENNSYLVANIA.  
XX  
PI Cooperman BS, Hirschmann RF, Laub P, Sascho S, Sprengeler PA;  
PI Barwis BA, Nair S, Fisher A, Smith AB;  
XX  
XX  
DR WPI; 2000-223191/19.  
XX  
XX  
PT Composition comprising an inhibitor of ribonucleotide reductase used in  
PT the treatment of disorders associated with DNA replication, such as  
PT tumor growth and metastasis and infectious disease -  
XX  
XX  
PS Disclosure; Column 4; 92pp; English.

CC The invention relates to a composition comprising a peptide inhibitor  
CC of ribonucleotide reductase. Peptides AAV83838-Y83864 represent examples  
CC of the inhibitor peptides used in the composition. The peptides are  
CC synthesised based on the sequences of known mammalian, bacterial and  
CC viral ribonucleotide reductases. The compositions are used in the  
CC treatment of disorders associated with DNA replication, such as tumour  
CC growth and metastasis and infectious disease, such as those caused by  
CC viruses, bacteria, fungi and parasites.  
XX  
XX  
SQ Sequence 7 AA;

Query Match 100.0%; Score 18; DB 21; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|  
|  
|  
|  
DB 2 SLDA 5

## RESULT 6

AAE05612  
ID AAE05612 standard; peptide; 7 AA.  
XX  
XX  
AC AAE05612;  
XX  
XX  
DT 24-SEP-2001 (first entry)  
DE N-terminal CT27 peptide obtained by proteolysis of Mt 45 kDa antigen.  
XX  
XX  
KW Mycobacterial disease; infection; 88-kDa protein; tuberculosis; TB;  
KW human immunodeficiency virus; HIV; CT27 glycopeptide; serodiagnosis;

```

KM fast atom bombardment-mass spectrometry; FAB-MS; proteolysis;
KW 45 kDa glycoprotein; MPT 32; chymotrypsin; trypsin.
XX
OS Mycobacterium tuberculosis.
XX
PN US6245331-B1.
XX
PD 12-JUN-2001.
XX
PF 31-DEC-1997; 97US-0001984.
XX
PR 02-JAN-1997; 97US-0034003.
XX
PA (UYNT-) UNIV NEW YORK MEDICAL CENT.
XX (COLS ) UNIV COLORADO STATE.
XX
PI Laal S, Zolla-Pazner S, Belisle JT;
XX WPI; 2001-424324/45.
XX
DR
XX
PT Detecting a mycobacterial disease (tuberculosis) in individuals
PT comprise assaying a biological sample for the presence of
PT anti-Mycobacterium tuberculosis antibodies or M. tuberculosis
PT antigen-antibody complex
XX
PS Example 3; Column 46; 96pp; English.
XX
CC The present invention relates to a method for early detection of active
CC mycobacterial disease or infection comprising assaying a biological fluid
CC sample for the presence of early antibodies specific for an 88-kDa
CC Mycobacterium tuberculosis (Mt) protein or immune complexes consisting of
CC an 88-kDa M. tuberculosis protein antigen complexed with an antibody
CC specific for the antigen. The method is useful for the early and rapid
CC detection of mycobacterial disease, particularly tuberculosis, in
CC individuals at heightened risk of developing tuberculosis (TB). This
CC or other immunocompromised individuals. The method is a rapid and
CC inexpensive screening procedure for detecting mycobacterial disease.
CC The present sequence is the N-terminal amino acid sequence of C127
CC glycoprotein. The N-terminal peptide is identified by fast atom
CC bombardment-mass spectrometry (FAB-MS) and N-terminal amino acid
CC sequence analysis of peptides generated by proteolysis of
CC Mt 45 kDa/MPT 32 glycoprotein with chymotrypsin/trypsin. The Mt 45
CC kDa/MPT 32 early antigen is used in the serodiagnosis of tuberculosis.
XX
SQ Sequence 7 AA:
Query Match 100.0%; Score 18; DB 22; Length 7;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLDA 4
Db 4 SLDA 7

```

```

PD 20-MAR-1997.
XX
PF 10-SEP-1996; 96MO-US14600.
XX
PR 11-SEP-1995; 95US-0526710.
XX 11-SEP-1995; 95US-0526708.
XX
PA (LJOL-) LA JOLLA CANCER RES FOUND.
XX
PI Pasqualini R, Ruoslahti E;
XX WPI; 1997-202359/18.
XX
DR
XX
PT Obtaining compound that homes to selected organ or tissue - by in
PT vivo panning method, specifically to identify brain, kidney,
PT angiogenic vasculature or tumour tissue homing peptide(s)
XX
PS Example 2; Page 53; 75pp; English.
XX
CC This synthetic peptide is an example of a brain-homing peptide
CC that was identified using a claimed method for obtaining
CC molecules that home to a selected organ or tissue. This in vivo
CC panning method typically involves administering a phage display
CC library to a subject, and identifying expressed peptides which
CC home to the desired organ or tissue, e.g. brain, kidney, angiogenic
CC vascular tissue or tumour tissue. The isolated peptides (see
CC AAW13412-52, AAW1181-86) can be used to target e.g. drugs, toxins or
CC labels to the selected organ/tissue (claimed) or to identify and/or
CC isolate target molecules (claimed). The peptides can be directly
CC identified in vivo, as compared to prior art in vitro screening
CC methods, which require further examination to see if they maintain
XX specificity in vivo.
XX
SQ Sequence 8 AA:
Query Match 100.0%; Score 18; DB 18; Length 8;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLDA 4
Db 4 SLDA 7

```

```

RESULT 7
AAW13438
ID AAW13438 standard; Peptide: 8 AA.
XX
AC AAW13438;
XX
DT 15-JAN-1998 (first entry)
XX
DE Brain homing peptide.
XX
KW Brain homing peptide; in vivo panning; screening; phage display;
XX drug delivery.
XX
OS Synthetic.
XX
PN MO9710507-A1.
XX

```

```

RESULT 8
AAB07398
ID AAB07398 standard; peptide: 8 AA.
XX
AC AAB07398;
XX
DT 17-OCT-2000 (first entry)
XX
DE Brain homing peptide # 12.
XX
KW Brain; homing peptide; organ targeting; tissue targeting; mouse; cyclic.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT Disulfide-bond 1..8
FT /note="Can optionally form a cyclic peptide"
XX
PN US6068829-A.
XX
PD 30-MAY-2000.
XX
PF 23-JUN-1997; 97US-0862855.
XX
PR 11-SEP-1995; 95US-0526710.
XX 10-MAR-1997; 97US-0813273.
XX
PA (BURN-) BURNHAM INST.
XX

```



PI Paqualini R, Ruoslahti E;  
XX  
DR WPI; 2000-410850/35.  
XX  
PT Identifying and recovering organ homing molecules or peptides by in  
XX vivo panning comprises administering a library of diverse peptides  
XX linked to a tag which facilitates recovery of these peptides -  
PS Example 2; Column 17; 20pp; English.  
XX  
CC The present sequence is a mouse brain homing peptide. This sequence was  
CC identified by using in vivo panning to screen a library of potential  
CC organ homing molecules. The present sequence can be used to direct a  
CC moiety to a the brain tissue, by linking the moiety to the present  
CC sequence. Examples of potential moieties are drugs, toxins or a  
CC detectable label.  
XX  
SQ Sequence 8 AA;  
Query Match 100.0%; Score 18; DB 21; Length 8;  
Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SLDA 4  
DB 4 SLDA 7  
RESULT 9  
AAV56002  
ID AAV56002 standard; peptide; 8 AA.  
XX  
AC AAV56002;  
XX  
DT 15-MAR-2000 (first entry)  
DE  
XX Mucor hlemalis endo-beta-N-acetylglucosaminidase peptide p60-AP-3.  
DE  
XX Endo-beta-N-acetylglucosaminidase; enzyme; cleavage; sugar chain;  
KW carboxylate; glycoprotein; disease; cancer.  
XX  
XX Mucor hlemalis.  
OS  
XX WO961591-A1.  
PN  
XX 02-DEC-1999.  
PD  
XX 20-MAY-1999; 99WO-JP02644.  
PF  
XX 22-MAY-1998; 98JP-0141717.  
PR  
XX (KIRI ) KIRIN BEER KK.  
PA  
XX Kobayashi K, Takeuchi M, Iwamatsu A, Yamamoto K, Kumagai H;  
PI Yoshida S;  
XX  
XX WPI: 2000-062704/05.  
DR  
XX Polynucleotides encoding endo-beta-N-acetylglucosaminidase, useful for  
PT the specific cleavage of complex sugar chains -  
XX  
XX Example 4; Page 22; 72pp; Japanese.  
CC Peptides AAV5594-Y56014 represent fragments of the Mucor hlemalis  
CC endo-beta-N-acetylglucosaminidase protein derived by cleavage with  
CC Lys-C protease and subsequent amino acid sequencing. The peptide  
CC sequences were then used to generate primers to isolate the coding  
CC sequence for isolating the corresponding gene. The protein is useful  
CC for the specific cleavage of complex sugar chains for the preparation,  
CC investigation and modification of these chains, especially in complex  
CC carbohydrates forming part of a glycoprotein molecule which undergoes  
CC modification in diseases such as cancer.

SQ Sequence 8 AA;  
Query Match 100.0%; Score 18; DB 21; Length 8;  
Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SLDA 4  
DB 3 SLDA 6  
RESULT 10  
AAU68419  
ID AAU68419 standard; Peptide; 8 AA.  
XX  
AC AAU68419;  
XX  
DT 16-JAN-2002 (first entry)  
DE  
XX Human Breast cancer-associated protein isoform, BPI-247 peptide #3.  
DE  
XX Human; Breast cancer-associated protein isoform; breast cancer;  
KW immunogen; cytostatic; BPI; tryptic digest peptide.  
XX  
XX Homo sapiens.  
OS  
XX WO200171357-A2.  
PN  
XX 27-SEP-2001.  
PD  
XX 20-MAR-2001; 2001WO-GB01219.  
PF  
XX 20-MAR-2000; 2000GB-0006695.  
PR  
XX 24-MAR-2000; 2000GB-0007265.  
XX  
XX (OXFO-) OXFORD GLYCOSCIENCES UK LTD.  
PA  
XX Herath HMA, O'Hare MJ, Page MJ, Parekh RB, Waterfield MD;  
PI WPI; 2001-611532/70.  
DR  
XX Identifying proteins for clinical screening, diagnosis and prognosis of  
PT breast cancer, comprises detecting Breast Cancer-Associated Protein  
PT Isoforms (BPis) using two-dimensional electrophoresis -  
XX  
XX Claim 9; Page 51; 197pp; English.  
PS  
XX The invention relates to diagnosing, determining the stage or severity,  
CC or identifying the risk of a subject developing cancer (especially  
CC breast cancer), or monitoring the effect of therapy on a subject with  
CC cancer, comprising analysing a test sample using two-dimensional  
CC electrophoresis and detecting Breast Cancer-Associated Protein  
CC Isoforms (BPis). The methods disclosed are used for the diagnosis and  
CC prognosis of breast cancer, for determining the severity of breast  
CC cancer, and for identifying a subject at risk of developing breast  
CC cancer, and monitoring the effect of therapy administered to a subject.  
CC Antibodies raised against the binding domain of a BPI, the binding domain  
CC of a BPI, a nucleic acid encoding a BPI, or a nucleic acid that inhibits  
CC the function of a BPI can be incorporated into a pharmaceutical  
CC composition for treating or preventing breast cancer. The methods use  
CC sensitive and specific biomarkers provide early diagnosis of breast  
CC cancer, and the compositions are more potent, specific, and has a more  
CC rapid effect with fewer side effects than other prior art methods.  
CC The present sequence is a tryptic digest peptide from a BPI of the  
CC invention.  
XX  
SQ Sequence 8 AA;  
Query Match 100.0%; Score 18; DB 22; Length 8;  
Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SLDA 4

DB 1 SUDA 4

## RESULT 11

ID AAE11804 standard; peptide; 8 AA.

AC AAE11804;

DT 18-DEC-2001 (first entry)

DE Phage peptide #12 targeted to brain.

KW Enriched library fraction; brain; kidney; tumour; panning; diagnostic;  
KW molecular medicine; drug delivery; peptidomimetic; pharmaceutical.

OS Bacteriophage.

PN US6296832-B1.

PD 02-OCT-2001.

PF 08-JAN-1999; 99US-0226985.

PR 23-JUN-1997; 97US-0862855.

PR 11-SEP-1995; 95US-0526710.

PR 10-MAR-1997; 97US-0813273.

PS (BURN-) BURNNHAM INSTR.

PI Ruoslahti E, Pasqualini R;

DR WPI; 2001-610691/70.

XX Enriched library fraction comprising molecules recovered by in vivo  
XX panning that selectively home to a selected organ or tissue useful for  
XX treating disease or in diagnostic methods

PS Example 2; Column 17; 21pp; English.

CC The invention relates to an enriched library fraction containing  
CC molecules that selectively home to a selected organ or tissue such as  
CC brain, kidney or tumour recovered by in vivo panning. The invention  
CC generally relates to the field of molecular medicine, drug delivery and  
CC to a method of in vivo panning for identifying a molecule that homes to a  
CC specific organ. The molecules, e.g., peptides, peptidomimetics, proteins  
CC and fragments of proteins contained in an enriched library fraction may  
CC be administered to a subject as part of a pharmaceutical composition to  
CC treat disease or in diagnostic methods. The present sequence is a  
CC peptide from bacteriophage targeted to brain.

SQ Sequence 8 AA;

Query Match 100.0%; Score 18; DB 22; Length 8;

Best Local Similarity 100.0%; Pred. No. 7.8e+05; Mismatches 0; Indels 0; Gaps 0;

OY 1 SUDA 4

DB 4 SUDA 7

## RESULT 12

ID ABG67883 standard; Peptide; 8 AA.

AC ABG67883;

DT 07-OCT-2002 (first entry)

DE Human ADPI tryptic digest peptide #592.

KW Human; Alzheimer's disease; AD; brain tissue; ADP; ADPI;  
KW Alzheimer's disease-associated feature; neuroprotective;  
KW Alzheimer's disease-associated protein isoform; nootropic;  
KW ADPI tryptic digest peptide.

OS Homo sapiens.

PN WO200246767-A2.

PD 13-JUN-2002.

PF 29-NOV-2001; 2001WO-GB05289.

PR 08-DEC-2000; 2000US-254431P.

PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.

PI Herath HMC, Parekh RB, Rohlf C;

DR WPI; 2002-508575/54.

XX Screening, diagnosis or prognosis of Alzheimer's disease in subject,  
XX comprises detecting Alzheimer's disease-associated features or  
XX Alzheimer's disease-associated protein isoforms in brain tissue  
XX from the subject

PS Claim 7; Page 65; 427pp; English.

CC The present invention relates to methods and compositions for the  
CC screening, diagnosis or prognosis of Alzheimer's disease (AD) in  
CC a subject. The method comprises analysing a sample of brain tissue  
CC from a subject by 2D electrophoresis to generate a 2D array of  
CC Alzheimer's disease-associated features (ADFs), whose relative  
CC abundance correlates with the presence, absence, stage or severity of  
CC AD and comparing the abundance of each feature with the abundance of  
CC that chosen feature in brain tissue from persons free from AD. The  
CC invention also describes Alzheimer's disease-associated protein  
CC isoforms (ADPIs) detectable in brain tissue. The methods and  
CC compositions of the invention are useful for the screening, diagnosis  
CC or prognosis of AD in a subject, for determining the stage or severity  
CC of AD in a subject, for identifying a subject at risk of developing AD,  
CC or for monitoring the effect of therapy administered to a subject  
CC having AD. Antibodies capable of binding to ADPIs are useful for  
CC treating or preventing AD, and for determining the efficacy of a given  
CC treatment regime. An agent that modulates the activity of ADPI is  
CC useful in the manufacture of a medicament for the treatment or  
CC prevention of AD in a subject. ABG67292-ABG68038 represent human ADPI  
CC tryptic digest peptides.

SQ Sequence 8 AA;

Query Match 100.0%; Score 18; DB 23; Length 8;

Best Local Similarity 100.0%; Pred. No. 7.8e+05; Mismatches 0; Indels 0; Gaps 0;

OY 1 SUDA 4

DB 1 SUDA 4

## RESULT 13

ID AAU10715 standard; peptide; 8 AA.

AC AAU10715;

DT 12-MAR-2002 (first entry)

DE Brain homing peptide #12 useful for delivery of target molecules.

KW Organ targeting; tissue targeting; cancer; tumour homing molecule;  
KW delivery of target molecule; brain homing peptide.

OS Synthetic.  
 XX US6306365-B1.  
 PN  
 XX  
 PD 23-OCT-2001.  
 XX  
 XX 08-JAN-1999; 99US-0227906.  
 PF  
 XX 23-JUN-1997; 97US-0862855.  
 PR 11-SEP-1995; 95US-0526710.  
 PR 10-MAR-1997; 97US-0813273.  
 XX  
 PA (BURN-) BURNHAM INST.  
 PI Ruostahli E, Pasqualini R;  
 XX WPI; 2002-040196/05.  
 DR  
 XX  
 PT Recovering molecules that home to an organ or tissue, useful for  
 PT identifying molecules that home to a specific organ or tissue, e.g.  
 PT identifying a tumour homing molecule to identify the presence of cancer,  
 PT by in vivo panning of a library -  
 PS Example 2; Column 17; 21pp; English.  
 XX  
 CC The present invention relates to a method of recovering molecules that  
 CC home to a selected organ or tissue. The method comprises administering  
 CC to the subject the library of diverse molecules, collecting a sample of  
 CC the selected organ or tissue (e.g. brain or kidney), and recovering from  
 CC the sample several molecules that home to the selected organ or tissue.  
 CC The method is useful for identifying molecules, particularly useful for  
 CC screening large number of molecules (e.g. peptides), that home to a  
 CC specific organ. The identified molecule is useful for e.g. raising an  
 CC antibody specific for a target molecule, targeting a desired moiety  
 CC (e.g. drug, toxin or detectable label) to the selected organ.  
 CC Specifically, the method is useful for identifying the presence of cancer  
 CC in a subject by linking an appropriate moiety to a tumour homing  
 CC molecule. The present method provides a direct means for identifying  
 CC molecules that specifically home to a selected organ and, therefore  
 CC provides a significant advantage over previous methods, which require  
 CC that a molecule identified using an in vitro screening method  
 CC subsequently be examined to determine if it maintains its specificity in  
 CC vivo. AAU10704-AAU10723 represent brain homing peptides described in  
 CC the present invention.  
 CC  
 SQ Sequence 8 AA;  
 QY 1 SIDA 4  
 DB 4 SIDA 7  
 IIII  
 RESULT 14  
 AAY93097  
 ID AAY93097 standard; peptide; 9 AA.  
 XX  
 AC AAY93097;  
 XX  
 DT 08-NOV-2000 (first entry)  
 XX  
 DE Transforming growth factor inhibitory peptide P143.  
 XX  
 KW Hepatotropic; antagonist; transforming growth factor beta1; TGF-beta1;  
 KW competitive inhibition; collagen synthesis stimulation inhibitor; liver;  
 KW extracellular matrix degradation inhibitor; mimotope; cirrhosis.  
 XX  
 OS Homo sapiens.  
 XX WO200031135-A1.  
 PN

XX 02-JUN-2000.  
 PD  
 XX 23-NOV-1999; 99WO-ES00375.  
 PF  
 XX  
 PR 24-NOV-1998; 98ES-0002465.  
 PA  
 XX (CIEN-) INST CIENTIFICO & TECNOLÓGICO NAVARRA.  
 PI Ezquerro Saenz IJ, Lasarte Sagastibelza JI, Prieto Valtuena J;  
 PI Borras Cuesta F;  
 XX WPI; 2000-411935/35.  
 DR  
 XX  
 PT Peptides that antagonize binding of transforming growth factor beta1,  
 PT useful for treatment of liver disease, especially cirrhosis, are  
 PT partial sequences of the factor or its receptors -  
 PS Disclosure; Page 31; 86pp; Spanish.  
 XX  
 CC The invention relates to synthetic peptides that antagonise the binding  
 CC of transforming growth (TGF) factor beta1 (TGF-beta1) to its receptor  
 CC in vivo which have partial amino acid sequences identical, or similar,  
 CC with those of TGF-beta1 and/or its receptors. Peptides AAY2945-Y93133  
 CC represent examples of the peptides of the invention. The peptides act  
 CC by competitive inhibition of the binding of TGF-beta1 to its receptors,  
 CC e.g. they are inhibitors of stimulation of collagen synthesis in liver  
 CC cells and inhibitors of synthesis of proteolytic enzymes able to degrade  
 CC the extracellular matrix. The peptides, their mimetopes and/or DNA (or  
 CC expression systems) encoding the peptides are used for treatment of  
 CC liver disease, specifically cirrhosis.  
 CC  
 SQ Sequence 9 AA;  
 QY 1 SIDA 4  
 DB 2 SIDA 5  
 IIII  
 RESULT 15  
 AAB99639  
 ID AAB99639 standard; peptide; 9 AA.  
 XX  
 AC AAB99639;  
 XX  
 DT 04-SEP-2001 (first entry)  
 XX  
 DE Human hsp60 derived peptide #65.  
 XX  
 KW Human; heat shock protein; hsp60; inhibition; antagonist; anti-allergic;  
 KW pro-inflammatory immune response; immune system; inflammatory disease;  
 KW immunostimulatory; anti-inflammatory; antidiabetic; immunosuppressive;  
 KW antirheumatic; nootropic; antiarteriosclerotic; antimicrobial; antidiabetic;  
 KW neuroprotective; autoimmune disease; diabetes; multiple sclerosis;  
 KW rheumatoid arthritis; myasthenia gravis; systemic lupus erythematosus;  
 KW chronic inflammation; atherosclerosis; inflammatory bowel disease;  
 KW reactive arthritis; chronic infection; graft rejection; gingivitis;  
 KW gastritis; ulcer; thrombosis; allergy; acute infection.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX WO200143691-A2.  
 PN  
 XX 21-JUN-2001.  
 PD  
 XX 14-DEC-2000; 2000WO-IL00833.  
 PF  
 XX 15-DEC-1999; 99EP-0125035.  
 PR

```

PR 15-DEC-1999; 9905-0172297.
XX
XX (PEPT-) PEPTOR LTD.
XX
XX Kolb H, Elias D;
XX
XX WPI; 2001-390150/41.
DR
XX
XX Anti-inflammatory peptides of hsp60, which are capable of acting as
XX PT antagonists of hsp60, useful for treating an inflammatory disease or
XX PT disorder, e.g. an autoimmune disease, diabetes, multiple sclerosis and
XX PT rheumatoid arthritis -
XX
XX Example 9; Page 58; 76pp; English.
XX
XX The present invention describes anti-inflammatory peptides of heat shock
XX CC protein hsp60, which are capable of acting as antagonists of hsp60
XX CC characterised by their ability to reduce or prevent the induction of a
XX CC pro-inflammatory immune response of cells of the innate immune system by
XX CC hsp60. The peptides have anti-inflammatory, antidiabetic, antirheumatic,
XX CC immunosuppressive, nootropic, antiarteriosclerotic, antimicrobial,
XX CC antiviral, antiallergic and neuroprotective activities. The antagonists
XX CC are useful in the treatment of an inflammatory disease or disorder,
XX CC e.g. an autoimmune disease, diabetes, multiple sclerosis, rheumatoid
XX CC arthritis, myasthenia gravis, systemic lupus erythematosus, chronic
XX CC inflammation, atherosclerosis, inflammatory bowel disease, reactive
XX CC arthritis, chronic infections, graft rejection, gingivitis, gastritis,
XX CC ulcer, thrombosis, allergy and acute infections. AAH43896 to AAH43899
XX CC and AAB9942 to AAB99673 represent sequences used in the exemplification
XX CC of the present invention.
XX
XX Sequence 9 AA:
SQ
Query Match 100.0%; Score 18; DB 22; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SLDA 4
| | | |
Db 2 SLDA 5
RESULT 16
AAR72570
ID AAR72570 standard; Peptide: 10 AA.
XX
XX AAR72570;
AC
XX
XX 23-NOV-1995 (first entry)
DT
XX
XX Penicillium purpurogenum endo-inulinase peptide fragment 175-184.
DE
XX
XX Penicillium purpurogenum variety rubrisclerotium; endo-inulinase;
XX KW inulo-oligosaccharides; inulin; peptide fragment 175-184.
XX
XX Penicillium purpurogenum.
OS
XX
XX JP07059574-A.
XX PN
XX
XX 07-MAR-1995.
XX PD
XX
XX 23-AUG-1993; 93JP-0229448.
XX PF
XX
XX 23-AUG-1993; 93JP-0229448.
XX PR
XX
XX (MITK ) MITSUI TOATSU CHEM INC.
XX PA
XX
XX WPI; 1995-135898/18.
XX DR
XX
XX A new DNA encoding endo-inulinase - useful for the production of
XX PT inulo-oligosaccharides
XX PS
XX Disclosure; Page 7; 8pp; Japanese.

```

```

XX
XX AAR72567-R72575 are peptide fragments of AAR72566, the Penicillium
XX CC purpurogenum variety rubrisclerotium endo-inulinase. The endo-inulinase
XX CC can be used to prepare an inulo-oligosaccharide from inulin.
XX
XX Sequence 10 AA:
SQ
Query Match. 100.0%; Score 18; DB 16; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SLDA 4
| | | |
Db 6 SLDA 9
RESULT 17
AAM61058
ID AAM61058 standard; peptide: 10 AA.
XX
XX AAM61058;
AC
XX
XX 25-SEP-1998 (first entry)
DT
XX
XX Peptide from IgY used in the course of the invention.
DE
XX
XX Immunoglobulin G; IgG; FcRn binding; half-life; increase; decrease;
XX KW treatment; idiopathic thrombocytopenic purpura; ITP;
XX KW Kawasaki disease; AIDS; Guillain-Barre syndrome; dermatomyositis.
XX
XX Synthetic.
XX OS
XX
XX WO9823289-A1.
XX PN
XX
XX 04-JUN-1998.
XX PD
XX
XX 26-NOV-1997; 97WO-US21437.
XX PF
XX
XX 27-NOV-1996; 96US-0031607.
XX PR
XX
XX (GEHO ) GEN HOSPITAL CORP.
XX PA
XX
XX (UYBR-) UNIV BRANDEIS.
XX PI
XX
XX Israel EJ, Simlster NE;
XX DR
XX
XX WPI; 1998-322461/28.
XX
XX New antibody having altered Fc receptor binding site - allows
XX PT increased or decreased half-lives and clearance rates, useful for,
XX PT e.g. tumour imaging or treatment of AIDS
XX
XX Disclosure; Fig 2B; 32pp; English.
XX PS
XX
XX AAM61058-60 represent partial amino acid sequence of different
XX CC immunoglobulin (Ig) molecules. The specification describes a
XX CC non-naturally occurring IgG having an altered amino acid sequence
XX CC compared to native IgG, which binds to FcRn with increased or
XX CC decreased affinity. FcRn is a receptor found on the intestinal surface
XX CC of the neonate and is responsible passage of maternal milk IgG from the
XX CC epithelial lumen to the systemic circulation via the intestinal
XX CC epithelial cells. FcRn is also used to prevent clearance of circulating
XX CC IgG from the circulation. Altering the binding site on an IgG for FcRn
XX CC allows greater or lower affinity for the FcRn receptor, which
XX CC subsequently increases or decreases the half-life of circulating IgG,
XX CC respectively. IgG's with an increased half-life can be used to treat
XX CC immune conditions or diseases such as idiopathic thrombocytopenic
XX CC purpura (ITP), Kawasaki disease, AIDS, Guillain-Barre syndrome and
XX CC dermatomyositis.
XX
XX Sequence 10 AA:
SQ
Query Match 100.0%; Score 18; DB 19; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;

```

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
| | | |  
Db 7 SLDA 10

## RESULT 18

AA686956  
ID AAG86956 standard; Peptide: 10 AA.

AC AAG86956;

DT 11-SEP-2001 (first entry)

DE Saccharomyces cerevisiae peptide, SEQ ID NO: 1905.

KM Saccharomyces cerevisiae; complementary peptide; peptide identification;  
KW drug discovery; drug design.

OS Saccharomyces cerevisiae.

PN WO200142276-A1.

PD 14-JUN-2001.

PF 13-DEC-2000; 2000WO-GB04773.

PR 13-DEC-1999; 99GB-0029471.

PA (PROT-) PROTEOM LTD.

PI Roberts GW, Heal JR;

DR WPI; 2001-367863/38.

PT Identifying complementary peptides by analysis of protein and  
nucleotide sequence databases, useful in drug design -

PS Example 3; Page 292; 488pp; English.

CC The invention relates to the identification of complementary peptides  
by analysis of protein and nucleotide sequence databases from higher  
eukaryotic genomes, excluding human and plants. The specific

CC complementary peptides interact with their relevant target proteins  
encoded in the eukaryotic genome. The peptides may be used as reagents  
and drugs for drug discovery and as lead ligands for drug design and

CC development. The present sequence is a complementary peptide from  
Saccharomyces cerevisiae.

CC Sequence 10 AA;

Query Match 100.0%; Score 18; DB 22; Length 10;  
Best Local Similarity 100.0%; Pred. NO. 1.5e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
| | | |  
Db 6 SLDA 9

## RESULT 19

AA92947  
ID AAY92947 standard; peptide: 12 AA.

AC AAY92947;

DT 08-NOV-2000 (first entry)

DE Transforming growth factor inhibitory peptide #3.

KM Hepatotropic; antagonist; transforming growth factor betal; TGF-b1;  
KW competitive inhibition; collagen synthesis stimulation inhibitor; liver;

KM extracellular matrix degradation inhibitor; mimotope; cirrhosis.

OS Homo sapiens.

PN WO200031135-A1.

PD 02-JUN-2000.

PF 23-NOV-1999; 99WO-ES00375.

PR 24-NOV-1998; 98ES-0002465.

PA (CIEN-) INST CIENTIFICO & TECNOLOGICO NAVARRA.

PI Ezquerro Saenz IJ, Lasarte Sagastibelza JJ, Prieto Valtuena J;

DR Borras Cuesta F;

WPI; 2000-411935/35.

PT Peptides that antagonize binding of transforming growth factor betal,  
useful for treatment of liver disease, especially cirrhosis, are

PT partial sequences of the factor or its receptors -

PS Claim 4; Page 80; 86pp; Spanish.

CC The invention relates to synthetic peptides that antagonise the binding  
of transforming growth (TGF) factor betal (TGF-b1) to its receptor

CC in vivo which have partial amino acid sequences identical, or similar,  
with those of TGF-b1 and/or its receptors. Peptides AAY92945-Y93133

CC represent examples of the peptides of the invention. The peptides act  
by competitive inhibition of the binding of TGF-b1 to its receptors,

CC e.g. they are inhibitors of stimulation of collagen synthesis in liver  
cells and inhibitors of synthesis of proteolytic enzymes able to degrade

CC the extracellular matrix. The peptides, their mimetopes and/or DNA (or  
expression systems) encoding the peptides are used for treatment of

CC liver disease, specifically cirrhosis.

CC Sequence 12 AA;

Query Match 100.0%; Score 18; DB 21; Length 12;  
Best Local Similarity 100.0%; Pred. NO. 1.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
| | | |  
Db 2 SLDA 5

## RESULT 20

AA93007  
ID AAY93007 standard; peptide: 12 AA.

AC AAY93007;

DT 08-NOV-2000 (first entry)

DE Transforming growth factor inhibitory peptide p53.

KM Hepatotropic; antagonist; transforming growth factor betal; TGF-b1;  
KW competitive inhibition; collagen synthesis stimulation inhibitor; liver;

KW extracellular matrix degradation inhibitor; mimotope; cirrhosis.

OS Rattus sp.

PN WO200031135-A1.

PD 02-JUN-2000.

PF 23-NOV-1999; 99WO-ES00375.

PR 24-NOV-1998; 98ES-0002465.

PA (CIEN-) INST CIENTIFICO & TECNOLOGICO NAVARRA.

XX Ezguero Saenz JI, Lasarte Sagastibelza JU, Prieto Valtuena J;  
PI Borras Cuesta F;  
XX WPI: 2000-411935/35.  
XX  
PT Peptides that antagonize binding of transforming growth factor betal,  
PT useful for treatment of liver disease, especially cirrhosis, are  
XX partial sequences of the factor or its receptors  
PS  
XX Disclosure: Page 27; 86pp; Spanish.  
CC  
CC The invention relates to synthetic peptides that antagonise the binding  
CC of transforming growth (TGF) factor betal (TGF-b1) to its receptor  
CC in vivo which have partial amino acid sequences identical, or similar,  
CC with those of TGF-b1 and/or its receptors. Peptides AAY92945-Y93133  
CC represent examples of the peptides of the invention. The peptides act  
CC by competitive inhibition of the binding of TGF-b1 to its receptors,  
CC e.g. they are inhibitors of stimulation of collagen synthesis in liver  
CC cells and inhibitors of synthesis of proteolytic enzymes able to degrade  
CC the extracellular matrix. The peptides, their mimetopes and/or DNA (or  
CC expression systems) encoding the peptides are used for treatment of  
CC liver disease, specifically cirrhosis.  
SQ  
Sequence 12 AA;  
Query Match 100.0%; Score 18; DB 21; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 SLDA 4  
DB 6 SLDA 9  
RESULT 21  
ID AAY93008 standard; peptide: 12 AA.  
XX  
AC AAY93008;  
XX  
DT 08-NOV-2000 (first entry)  
XX  
DE Transforming growth factor inhibitory peptide P54.  
XX  
DE Hepatotropic; antagonist; transforming growth factor betal; TGF-b1;  
KW competitive inhibition; collagen synthesis stimulation inhibitor; liver;  
KW extracellular matrix degradation inhibitor; mimetope; cirrhosis.  
XX  
OS Rattus sp.  
XX  
PN WO200031135-A1.  
XX  
PD 02-JUN-2000.  
XX  
PD 23-NOV-1999; 99WO-ES00375.  
XX  
PR 24-NOV-1998; 98ES-0002465.  
XX  
XX (CIEN-) INST CIENTIFICO & TECNOLOGICO NAVARRA.  
XX  
XX Ezguero Saenz JI, Lasarte Sagastibelza JU, Prieto Valtuena J;  
PI Borras Cuesta F;  
XX  
XX WPI: 2000-411935/35.  
XX  
XX Peptides that antagonize binding of transforming growth factor betal,  
XX useful for treatment of liver disease, especially cirrhosis, are  
XX partial sequences of the factor or its receptors  
PS  
XX Disclosure: Page 27; 86pp; Spanish.  
CC  
CC The invention relates to synthetic peptides that antagonise the binding

CC of transforming growth (TGF) factor betal (TGF-b1) to its receptor  
CC in vivo which have partial amino acid sequences identical, or similar,  
CC with those of TGF-b1 and/or its receptors. Peptides AAY92945-Y93133  
CC represent examples of the peptides of the invention. The peptides act  
CC by competitive inhibition of the binding of TGF-b1 to its receptors,  
CC e.g. they are inhibitors of stimulation of collagen synthesis in liver  
CC cells and inhibitors of synthesis of proteolytic enzymes able to degrade  
CC the extracellular matrix. The peptides, their mimetopes and/or DNA (or  
CC expression systems) encoding the peptides are used for treatment of  
CC liver disease, specifically cirrhosis.  
SQ  
Sequence 12 AA;  
Query Match 100.0%; Score 18; DB 21; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 SLDA 4  
DB 2 SLDA 5  
RESULT 22  
ID AAY93009 standard; peptide: 12 AA.  
XX  
AC AAY93009;  
XX  
DT 08-NOV-2000 (first entry)  
XX  
DE Transforming growth factor inhibitory peptide P55.  
XX  
DE Hepatotropic; antagonist; transforming growth factor betal; TGF-b1;  
KW competitive inhibition; collagen synthesis stimulation inhibitor; liver;  
KW extracellular matrix degradation inhibitor; mimetope; cirrhosis.  
XX  
OS Rattus sp.  
XX  
PN WO200031135-A1.  
XX  
PD 02-JUN-2000.  
XX  
PD 23-NOV-1999; 99WO-ES00375.  
XX  
PR 24-NOV-1998; 98ES-0002465.  
XX  
XX (CIEN-) INST CIENTIFICO & TECNOLOGICO NAVARRA.  
XX  
XX Ezguero Saenz JI, Lasarte Sagastibelza JU, Prieto Valtuena J;  
PI Borras Cuesta F;  
XX  
XX WPI: 2000-411935/35.  
XX  
XX Peptides that antagonize binding of transforming growth factor betal,  
XX useful for treatment of liver disease, especially cirrhosis, are  
XX partial sequences of the factor or its receptors  
PS  
XX Disclosure: Page 27; 86pp; Spanish.  
XX  
XX The invention relates to synthetic peptides that antagonise the binding  
XX of transforming growth (TGF) factor betal (TGF-b1) to its receptor  
XX in vivo which have partial amino acid sequences identical, or similar,  
XX with those of TGF-b1 and/or its receptors. Peptides AAY92945-Y93133  
XX represent examples of the peptides of the invention. The peptides act  
XX by competitive inhibition of the binding of TGF-b1 to its receptors,  
XX e.g. they are inhibitors of stimulation of collagen synthesis in liver  
XX cells and inhibitors of synthesis of proteolytic enzymes able to degrade  
XX the extracellular matrix. The peptides, their mimetopes and/or DNA (or  
XX expression systems) encoding the peptides are used for treatment of  
XX liver disease, specifically cirrhosis.  
SQ  
Sequence 12 AA;

Query Match 100.0%; Score 18; DB 21; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
 ||||  
 Db 1 SLDA 4

RESULT 23  
 AAY93093

ID AAY93093 standard; peptide: 12 AA.

XX AAY93093;

XX 08-NOV-2000 (first entry)

XX Transforming growth factor inhibitory peptide p139.

XX Hepatotropic; antagonist; transforming growth factor betaf: TGF- $\beta$ 1;

KM competitive inhibition; collagen synthesis stimulation inhibitor; liver;

XX extracellular matrix degradation inhibitor; mimetope; cirrhosis.

OS Homo sapiens.

XX MO200031135-A1.

XX 02-JUN-2000.

XX 23-NOV-1999; 99NO-ES00375.

XX 24-NOV-1998; 98ES-0002465.

XX (CIEN-) INST CIENTIFICO & TECNOLÓGICO NAVARRA.

PI Esquerro Saenz J, Lasarte Sagastibelza JJ, Prieto Valtuena J;

PI Borras Cuesta F;

XX WPI; 2000-411935/35.

PT Peptides that antagonize binding of transforming growth factor betaf,

XX useful for treatment of liver disease, especially cirrhosis, are

XX partial sequences of the factor or its receptors -

PS Disclosure; Page 31; 86pp; Spanish.

XX The invention relates to synthetic peptides that antagonise the binding

CC of transforming growth (TGF) factor betaf (TGF- $\beta$ 1) to its receptor

CC in vivo which have partial amino acid sequences identical, or similar,

CC with those of TGF- $\beta$ 1 and/or its receptors. Peptides AAY92945-Y93133

CC represent examples of the peptides of the invention. The peptides act

CC by competitive inhibition of the binding of TGF- $\beta$ 1 to its receptors,

CC e.g. they are inhibitors of stimulation of collagen synthesis in liver

CC cells and inhibitors of synthesis of proteolytic enzymes able to degrade

CC the extracellular matrix. The peptides, their mimetopes and/or DNA (or

CC expression systems) encoding the peptides are used for treatment of

CC liver disease, specifically cirrhosis.

XX Sequence 12 AA:

Query Match 100.0%; Score 18; DB 21; Length 12;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
 ||||  
 Db 2 SLDA 5

RESULT 24  
 AAW69624

ID AAW69624 standard; peptide: 13 AA.

AC AAW69624;  
 XX 19-OCT-1998 (first entry)

DE Ste2 agonist from the alpha-Mid-5 library #4.

KM Yeast; Saccharomyces cerevisiae; pheromone; alpha factor; receptor;

XX surrogate; screening; selection.

XX Synthetic.

OS Saccharomyces cerevisiae.

XX US5789184-A.

XX 04-AUG-1998.

XX 05-JUN-1995; 95US-0464531.

XX 05-JUN-1995; 95US-0464531.

XX 31-MAR-1993; 93US-0041431.

XX 31-JAN-1994; 94US-0190328.

XX 20-SEP-1994; 94US-0309313.

XX 13-OCT-1994; 94US-0322137.

XX (CADU-) CADUS PHARM CORP.

XX Broach J, Fowlkes DM, Klein C, Manfredi J, Murphy AJ;

XX Paul J, Trueheart J;

XX WPI; 1998-446076/38.

XX N-PSDB; AAV40413.

XX Example 4; Column 143; 93pp; English.

XX The present invention describes a yeast cell having a pheromone system,

CC in which the cell comprises: (a) a first heterologous gene encoding a

CC heterologous surrogate of a yeast pheromone system protein, the surrogate

CC being a kinase and performing in the pheromone system of the yeast cell

CC a function naturally performed by the corresponding yeast pheromone

CC system protein; and (b) a second heterologous gene encoding a

CC heterologous peptide, where the heterologous peptide modulates the

CC interaction of the surrogate with the pheromone system in the yeast cell,

CC and the modulation is a selectable or screenable event. The yeast cells

CC are used in assaying a peptide for modulation of the activity of a non-

CC yeast surrogate for a pheromone system protein and determining by

CC detecting a change in the selectable or screenable event whether the

CC pheromone signal pathway is activated or inhibited by the interaction of

CC the surrogate and the peptide. The present sequence represents an

CC Ste2 agonist from the alpha-Mid-5 library which is from an example

XX of the present invention.

XX Sequence 13 AA:

Query Match 100.0%; Score 18; DB 19; Length 13;

Best Local Similarity 100.0%; Pred. No. 2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
 ||||  
 Db 5 SLDA 8

RESULT 25  
 AAW74421

ID AAW74421 standard; peptide: 13 AA.

XX AAW74421;

XX 10-MAY-1999 (first entry)

DE Ste2 agonist M2 peptide sequence.  
 XX  
 KW Yeast pheromone; Ste2 agonist; cognate yeast pheromone system protein;  
 KW farnesyl transferase; anticancer therapy.  
 XX  
 OS Synthetic.  
 XX  
 PN US5876951-A.  
 XX  
 PD 02-MAR-1999.  
 XX  
 PE 05-JUN-1995; 94US-0461598.  
 XX  
 PR 05-JUN-1995; 95US-0461598.  
 PR 31-MAR-1993; 93US-0041431.  
 PR 31-JAN-1994; 94US-0190328.  
 PR 20-SEP-1994; 94US-0309313.  
 PR 13-OCT-1994; 94US-0322137.  
 XX  
 PA (CADU-) CADUS PHARM CORP.  
 XX  
 PI Broach J, Fowlkes DM, Klein C, Manfredi J, Murphy AJ;  
 PI Paul J, Trueheart J.  
 XX  
 DR WPI: 1999-189631/16.  
 DR N-PSDB: AAX18211.  
 XX  
 PT Yeast cells having an engineered pheromone system - useful for  
 PT identifying drugs which can inhibit or activate pheromone system  
 PT protein, e.g. to develop anti-cancer therapies  
 XX  
 PS Example 4; Column 56; 93pp; English.  
 XX  
 CC This sequence represents an Ste2 agonist peptide sequence.  
 CC The invention relates to yeast cells engineered to express an exogenous  
 CC protein capable of substituting for a yeast protein involved in the  
 CC post-translational modification, transport, recognition or signal  
 CC transduction of a yeast pheromone. The system can be used to identify  
 CC drugs which inhibit or activate the ability of the surrogate to  
 CC substitute for the cognate yeast pheromone system protein. Inhibitors of  
 CC farnesyl transferase identified can be used for anticancer therapies.  
 CC  
 XX  
 SQ Sequence 13 AA;  
 XX  
 QY Query Match 100.0%; Score 18; DB 20; Length 13;  
 QY Best Local Similarity 100.0%; Pred. No. 2e+02;  
 QY Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 DB 1 SLDA 4  
 DB 5 SLDA 8  
 XX  
 RESULT 26  
 AAB20756  
 ID AAB20756 standard; peptide; 13 AA.  
 XX  
 AC AAB20756;  
 XX  
 DT 21-DEC-2000 (first entry)  
 XX  
 DE Alpha-Mid-5 library Ste2 agonist peptide sequence SEQ ID NO:96.  
 XX  
 KW Yeast; pheromone; alpha-factor; transporter; pheromone receptor;  
 KW G alpha subunit; MF alpha 1; MFAL; STE2; STE3; Gsa receptor; GPAL;  
 KW G protein coupled receptor; mutagenesis; amplification; screening;  
 KW hybrid; agonist; antagonist; signal transduction; detection;  
 KW identification.  
 XX  
 OS Saccharomyces cerevisiae.  
 OS Synthetic.  
 XX  
 PN US6100042-A.

XX  
 PD 08-AUG-2000.  
 XX  
 PE 13-OCT-1994; 94US-0322137.  
 XX  
 PR 31-MAR-1993; 93US-0041431.  
 PR 31-JAN-1994; 94US-0190328.  
 PR 20-SEP-1994; 94US-0309313.  
 XX  
 PA (CADU-) CADUS PHARM CORP.  
 XX  
 PI Fowlkes DM, Broach J, Klein C, Manfredi J, Murphy AJ, Trueheart J;  
 PI Manfredi J;  
 XX  
 DR WPI: 2000-531665/48.  
 XX  
 PT Mixture of recombinant yeast cells comprising a heterologous G protein  
 PT coupled receptor whose signal transduction activity is modulated by a  
 PT heterologous polypeptide which provides a detectable signal on  
 PT modulation -  
 XX  
 PS Example 4; Column 57; 95pp; English.  
 XX  
 CC The present invention describes recombinant yeast cell mixtures (I).  
 CC Each (I) has a heterologous G protein coupled receptor (GPCR) expressed  
 CC in the cell membrane such that signal transduction (ST) activity via  
 CC GPCR is modulated by interaction of extracellular region (ER) of GPCR  
 CC with a heterologous polypeptide (P) which interacts with ER of receptor.  
 CC Modulation of the ST activity by (P) provides a detectable signal.  
 CC Also described is a recombinant yeast cell (II) that has a cell membrane  
 CC which comprises a GPCR such that ST activity via GPCR is modulated by  
 CC interaction of an ER of GPCR with an extracellular signal, and a (P)  
 CC (I) is used for identifying a modulator of (P) expressed by the yeast  
 CC cell which involves providing (I) which comprises heterologous GPCR and  
 CC a heterologous test polypeptide, allowing the cells within the mixture  
 CC to generate a detectable signal and then identifying the heterologous  
 CC test peptide as a modulator of the heterologous receptor protein  
 CC expressed by the yeast cell. The yeast cells may be used to identify  
 CC drugs which inhibit or activate, to a detectable degree, the ability of  
 CC the surrogate to substitute for the cognate yeast pheromone system  
 CC proteins. The yeast cell is also used to screen agonists and  
 CC antagonists. The present sequence is used in the exemplification of  
 CC the present invention.  
 XX  
 SQ Sequence 13 AA;  
 XX  
 QY Query Match 100.0%; Score 18; DB 21; Length 13;  
 QY Best Local Similarity 100.0%; Pred. No. 2e+02;  
 QY Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 DB 1 SLDA 4  
 DB 5 SLDA 8  
 XX  
 RESULT 27  
 AAG79149  
 ID AAG79149 standard; peptide; 13 AA.  
 XX  
 AC AAG79149;  
 XX  
 DT 03-JAN-2002 (first entry)  
 XX  
 DE Amino acid sequence of an alpha-factor pheromone agonist.  
 XX  
 KW Cellular receptor; ion channel; cellular activity; drug discovery;  
 KW orphan receptor ligand; alpha-factor pheromone.  
 XX  
 OS Synthetic.  
 OS  
 PN US2001026926-A1.  
 XX



```

PD 04-OCT-2001.
XX
XX 21-DEC-2000; 2000US-0747774.
XX
XX 17-JAN-1996; 960US-0582333.
XX
XX 31-MAR-1993; 930US-0041431.
XX
XX 31-JAN-1994; 940US-0190328.
XX
XX 20-SEP-1994; 940US-0309313.
XX
XX 13-OCT-1994; 940US-0322137.
XX
XX 05-JUN-1995; 950US-0461383.
XX
XX 05-JUN-1995; 950US-0461598.
XX
XX 05-JUN-1995; 950US-0463181.
XX
XX 05-JUN-1995; 950US-0463531.
XX
XX (CADU-) CADUS PHARM CORP.
XX
XX Klein CA, Murphy AJ, Fowlkes DM, Broach J, Manfredi J, Paul J;
XX Trueheart J;
XX
XX WPI: 2001-615870/71.
XX
XX N-PSDB; AA165730.
XX
XX Identification of compounds modulating cellular receptor activity
XX useful for identifying and screening for ligands for orphan receptors,
XX comprises using recombinant cells comprising both receptors and test
XX polypeptide -
XX
XX Example 4; Page 31; 50pp; English.
XX
XX The specification describes an assay for screening and identifying
XX pharmaceutically effective compounds that specifically interact with
XX and modulate the activity of a cellular receptor or ion channel. The
XX assay uses a mixture of recombinant cells, each comprising a receptor
XX protein whose signal transduction activity is modulated by an
XX interaction with an extracellular signal, a recombinant gene encoding
XX a potential receptor polypeptide, and a reporter gene construct. The
XX assay is useful for rapid screening of large numbers of polypeptides to
XX identify polypeptides antagonizing or agonizing receptor activity, and
XX to identify drugs for modulating cellular activity. It is especially
XX useful to identify ligands for orphan receptors, especially ligands for
XX orphan cell surface receptors, which are useful in drug discovery. The
XX present sequence represents an alpha-factor phenomene agonist, which
XX was identified using the assay of the invention.
XX
XX Sequence 13 AA;
XX
XX Query Match 100.0%; Score 18; DB 22; Length 13;
XX Best Local Similarity 100.0%; Pred. No. 2e+02;
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 SLDA 4
XX ||||
XX DB 5 SLDA 8
XX
XX RESULT 28
XX AAB84497
XX ID AAB84497 standard; peptide: 13 AA.
XX
XX AAB84497;
XX
XX 05-SEP-2001 (first entry)
XX
XX Amino acid sequence of a Ste2 agonist from alpha-Mid-5 library.
XX
XX G protein coupled receptor; GPCR; cellular receptor; ion channel;
XX surrogate ligand; orphan receptor; Ste2 agonist.
XX
XX Synthetic.
XX
XX US625059-B1.
XX
XX 03-JUL-2001.

```

```

XX
XX 17-JAN-1996; 960US-0582333.
XX
XX 31-MAR-1993; 930US-0041431.
XX
XX 31-JAN-1994; 940US-0190328.
XX
XX 20-SEP-1994; 940US-0309313.
XX
XX 13-OCT-1994; 940US-0322137.
XX
XX 05-JUN-1995; 950US-0463181.
XX
XX (CADU-) CADUS PHARM CORP.
XX
XX Klein CA, Murphy AJM, Fowlkes DM, Broach J, Manfredi J, Paul J;
XX Trueheart J;
XX
XX WPI: 2001-396879/42.
XX
XX N-PSDB; AAH27806.
XX
XX Identifying a ligand for an orphan G protein coupled receptor comprises
XX using an recombinant yeast expression library -
XX
XX Example 4; Column 57-58; 128pp; English.
XX
XX The specification describes a method for identifying a ligand for
XX an orphan G protein coupled receptor (GPCR). The method comprises
XX rapidly screening large numbers of polypeptides in a yeast expression
XX library to identify those polypeptides which induce or antagonise
XX receptor bioactivity. The method is useful for screening and identifying
XX pharmaceutically effective compounds that specifically interact with
XX and modulate the activity of a cellular receptor or ion channel. The
XX assay is particularly amenable for identifying surrogate ligands for
XX orphan receptors. The present sequence represents a Ste2 agonist,
XX identified from an alpha-Mid-5 library, using the method of the
XX invention.
XX
XX Sequence 13 AA;
XX
XX Query Match 100.0%; Score 18; DB 22; Length 13;
XX Best Local Similarity 100.0%; Pred. No. 2e+02;
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 SLDA 4
XX ||||
XX DB 5 SLDA 8
XX
XX RESULT 29
XX AAY92950
XX ID AAY92950 standard; peptide: 14 AA.
XX
XX AAY92950;
XX
XX 08-NOV-2000 (first entry)
XX
XX Transforming growth factor inhibitory peptide #6.
XX
XX Hepatotropic; antagonist; transforming growth factor betal; TGF- $\beta$ 1;
XX competitive inhibition; collagen synthesis stimulation inhibitor; liver;
XX extracellular matrix degradation inhibitor; mimeope; cirrhosis.
XX
XX Homo sapiens.
XX
XX WO200031135-A1.
XX
XX 02-JUN-2000.
XX
XX 23-NOV-1999; 99MO-ES00375.
XX
XX 24-NOV-1998; 98ES-0002465.
XX
XX (CIEN-) INST CIENTIFICO & TECNOLOGICO NAVARRA.
XX
XX Ezquerro Saenz JF, Lasarte Sagastibelza J, Prieto Valtuena J;
XX Borrás Cuesta F;
XX

```

XX	WP1; 2000-411935/35.
XX	
PT	Peptides that antagonize binding of transforming growth factor beta1,
PT	useful for treatment of liver disease, especially cirrhosis, are
PT	partial sequences of the factor or its receptors -
XX	
PS	Claim 7; Page 81; 86pp; Spanish.
CC	
CC	The invention relates to synthetic peptides that antagonise the binding
CC	of transforming growth (TGF) factor beta1 (TGF-beta1) to its receptor
CC	in vivo which have partial amino acid sequences identical, or similar,
CC	with those of TGF-beta1 and/or its receptors. Peptides AAY92945-Y93133
CC	represent examples of the peptides of the invention. The peptides act
CC	by competitive inhibition of the binding of TGF-beta1 to its receptors,
CC	e.g. they are inhibitors of stimulation of collagen synthesis in liver
CC	cells and inhibitors of synthesis of proteolytic enzymes able to degrade
CC	the extracellular matrix. The peptides, their mimetopes and/or DNA (or
CC	expression systems) encoding the peptides are used for treatment of
CC	liver disease, specifically cirrhosis.
XX	
SQ	Sequence 14 AA:
Query Match	100.0%; Score 18; DB 21; Length 14;
Best Local Similarity	100.0%; Pred. NO. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0.	
QY	1 SLDA 4 
Db	2 SLDA 5
RESULT 30	
ID	AAY93098
AC	AAY93098 standard; peptide; 14 AA.
XX	
AC	AAY93098;
XX	
DT	08-NOV-2000 (first entry)
XX	
DE	Transforming growth factor inhibitory peptide P144.
XX	
KW	Hepatotropic; antagonist; transforming growth factor beta1; TGF-beta1;
KW	competitive inhibition; collagen synthesis stimulation inhibitor; liver;
KW	extracellular matrix degradation inhibitor; minnetope; cirrhosis.
XX	
OS	Homo sapiens.
XX	
PN	WO200031135-A1.
XX	
PD	02-JUN-2000.
XX	
PF	23-NOV-1999; 99WO-ES00375.
XX	
R	24-NOV-1998; 98ES-0002465.
XX	
PA	(CIEN-) INST CIENTIFICO & TECNOLOGICO NAVARRA.
PI	Ezquerro Saenz IJ, Lasarte Sagastibelza JI, Prieto Valtuena J;
PI	Borrás Cuesta F;
XX	
DR	WP1; 2000-411935/35.
XX	
PT	Peptides that antagonize binding of transforming growth factor beta1,
PT	useful for treatment of liver disease, especially cirrhosis, are
PT	partial sequences of the factor or its receptors -
XX	
PS	Disclosure: Page 31; 86pp; Spanish.

CC	represent examples of the peptides of the invention. The peptides act
CC	by competitive inhibition of the binding of TGF- $\beta$ 1 to its receptors,
CC	e.g. they are inhibitors of stimulation of collagen synthesis in liver
CC	cells and inhibitors of synthesis of proteolytic enzymes able to degrade
CC	the extracellular matrix. The peptides, their mimetopes and/or DNA (or
CC	expression systems) encoding the peptides are used for treatment of
CC	liver disease, specifically cirrhosis.
XX	
XX	Sequence 14 AA:
SO	
	Query Match 100.0%; Score 18; DB 21; Length 14;
	Best Local Similarity 100.0%; Pred. NO. 2.2e+02;
	Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0
OY	1 SLDA 4
Db	2 SLDA 5
RESULT 31	
AAM97105	
ID	AAM97105 standard; Peptide; 14 AA.
XX	
AC	AAM97105;
XX	
DT	24-JAN-2002 (first entry)
DE	
DE	Human peptide #380 encoded by a SNP oligonucleotide.
XX	
KW	Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW	neuroprotective; antimicrobial; gene therapy; vaccine; amyase; cancer;
KW	amyloid protein; angiotensin; apoptosis related protein; cadherin;
KW	cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW	complement related protein; cytochrome; kinesin; cytokine; interferon;
KW	interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW	multifactorial disease; autoimmune disease; infection;
KW	nervous system disease.
XX	
OS	Homo sapiens.
XX	
PN	MO200147944-A2.
XX	
PD	05-JUL-2001.
XX	
PF	28-DEC-2000; 2000WO-US35498.
XX	
PR	28-DEC-1999; 99US-0173419.
XX	
PR	27-DEC-2000; 2000US-0173419.
XX	
PA	(CURA-) CURAGEN CORP.
XX	
PI	Shinkets RA, Leach M;
XX	
DR	WPI; 2001-465210/50.
XX	
PT	Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT	oncogenes and histones, useful for diagnosing and treating, e.g.
PT	cancer, autoimmune diseases and infections -
XX	
PS	Disclosure: Page 3751; 4143pp; English.
XX	
CC	The present invention relates to oligonucleotides (see AAL26793-AAL34659)
CC	encoding polymorphic variants of proteins related to amylases, amyloid
CC	proteins, angiotensin, apoptosis related proteins, cadherin, cyclin,
CC	polymerase, oncogenes, histones, kinases, colony stimulating factors,
CC	complement related proteins, cytochromes, kinesins, cytokines,
CC	interferons, interleukins, G-protein coupled receptors and thioesterases.
CC	The present sequence is a peptide encoded by one such oligonucleotide.
CC	The oligonucleotides and the peptides encoded by them may be used in the
CC	prevention, diagnosis and treatment of diseases associated with
CC	inappropriate expression of the proteins listed above. Disorders that may
CC	be prevented, diagnosed and/or treated include multifactorial diseases
CC	with a genetic component, such as autoimmune diseases (e.g. rheumatoid
CC	arthritis).

CC arthritis, multiple sclerosis, diabetes, systemic lupus erythromatosus  
CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,  
CC brain, breast, colon and kidney, leukaemia), diseases of the nervous  
CC system and an infection of pathogenic organisms.

XX Sequence 14 AA:

Query Match 100.0%; Score 18; DB 22; Length 14;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
|||||

DB 10 SLDA 13

RESULT 32  
AAM97480  
ID AAM97480 standard; Peptide; 14 AA.

XX AAM97480;

XX 24-JAN-2002 (first entry)

XX Human peptide #755 encoded by a SNP oligonucleotide.

XX Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;  
XX neutroprotective; antimicrobial; gene therapy; vaccine; amylose; cancer;  
XX amyloid protein; angiotensin; apoptosis related protein; cadherin;  
XX cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;  
XX complement related protein; cytochrome; kinesin; cytokine; interferon;  
XX interleukin; G-protein coupled receptor; thioesterase; inflammation;  
XX multifactorial disease; autoimmune disease; infection;  
XX nervous system disease.

XX Homo sapiens.

XX WO200147944-A2.

XX 05-JUL-2001.

XX 28-DEC-2000; 2000WO-US35498.

XX 28-DEC-1999; 99US-0173419.

XX 27-DEC-2000; 2000US-0173419.

XX (CURA-) CURAGEN CORP.

XX Shinkets RA, Leach M;

XX WPI: 2001-465210/50.

XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,  
XX oncogenes and histones, useful for diagnosing and treating, e.g.  
XX cancer, autoimmune diseases and infections -

XX Disclosure: Page 3833; 4143pp; English.

XX The present invention relates to oligonucleotides (see AAL26793-AAL34659)  
XX encoding polymorphic variants of proteins related to amylases, amyloid  
XX proteins, angiotensin, apoptosis related proteins, cadherin, cyclin,  
XX polymerase, oncogenes, histones, kinases, colony stimulating factors,  
XX complement related proteins, cytochromes, kinesins, cytokines,  
XX interferons, interleukins, G-protein coupled receptors and thioesterases.  
XX The present sequence is a peptide encoded by one such oligonucleotide.  
XX The oligonucleotides and the peptides encoded by them may be used in the  
XX prevention, diagnosis and treatment of diseases associated with  
XX inappropriate expression of the proteins listed above. Disorders that may  
XX be prevented, diagnosed and/or treated include multifactorial diseases  
XX with a genetic component, such as autoimmune diseases (e.g. rheumatoid  
XX arthritis, multiple sclerosis, diabetes, systemic lupus erythromatosus  
XX and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,  
XX brain, breast, colon and kidney, leukaemia), diseases of the nervous

CC system and an infection of pathogenic organisms.

XX Sequence 14 AA:

Query Match 100.0%; Score 18; DB 22; Length 14;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
|||||

DB 2 SLDA 5

RESULT 33  
AAR93618  
ID AAR93618 standard; peptide; 15 AA.

XX AAR93618;

XX 26-APR-1996 (first entry)

XX HIV principal neutralisation epitope binding to 2F5 antibody.

XX SPNE; selected principal neutralisation epitope; vaccine; HIV;  
XX outer membrane proteosome; Neisseria; OMPC; AIDS; 2F5 antibody.

XX Synthetic.

XX GB282379-A.

XX 05-APR-1995.

XX 23-SEP-1994; 94GB-0019255.

XX 30-SEP-1993; 93US-0129997.

XX (MERI) MERCK & CO INC.

XX Arnold BA, Conley AJ, Kessler JA;

XX WPI: 1995-125266/17.

XX New antigenic conjugate useful as vaccine for AIDS - comprising HIV  
XX principal neutralisation epitope covalently linked to outer membrane  
XX proteosome of Neisseria

XX Claim 14; Page 8; 75pp; English.

XX An antigenic conjugate, useful as a vaccine for AIDS, has the formula  
XX (SPNE)<sub>n</sub>(OMPC), where SPNE is a selected principal neutralisation  
XX epitope of HIV, which is one of 27 specified polypeptides (including the  
XX present sequence) or their fragments containing at least 5 amino acids  
XX and including the DRW or DKW region; OMPC is purified outer membrane  
XX proteosome of Neisseria (pref. N. meningitidis); and n is 1-200,  
XX indicating the number of SPNE moieties covalently linked to the OMPC.  
XX The conjugates may be substituted by anions, and conjugation  
XX may be via a biogenic spacer. The SPNE polypeptides bind an HIV broadly  
XX neutralising monoclonal antibody (2F5 antibody) specific for the  
XX ectodomain of HIV gp41 transmembrane glycoprotein. They were originally  
XX identified in the screening of phage epitope libraries having  
XX randomly generated epitope polypeptides accessible to the antibody.  
XX The library used was library Alpha described in AAR83295. The sequences  
XX of these polypeptides were deduced from their corresponding DNA  
XX sequence, determined by PCR.

XX Sequence 15 AA:

Query Match 100.0%; Score 18; DB 16; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
|||||

Db 7 SLDA 10

RESULT 34  
 AAY95420  
 ID AAY95420 standard; Peptide; 16 AA.

XX AAY95420;

XX 25-SEP-2000 (first entry)

XX Anti-angiogenic D3 peptide.

XX Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;

XX endothelial cell proliferation; apoptosis; cancer; ocular disorder;

XX rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;

XX therapy; human; D3 peptide.

XX Homo sapiens.

XX WO200035407-A2.

XX 22-JUN-2000.

XX 02-DEC-1999; 99WO-US28465.

XX 16-DEC-1998; 98US-0112427.

XX (UTEM ) UNIV TEMPLE.

XX (MCCR/) MCCR/ R K.

XX MCCR/ R K.

XX WPI; 2000-442247/38.

XX Composition for inhibiting angiogenesis and endothelial cell

XX proliferation, inducing endothelial cell apoptosis and treating cancer,

XX rheumatoid arthritis, and ocular disorders comprises a kininogen domain

XX 3 analog -

XX Claim 20; Page 29; 44pp; English.

XX The present sequence is that of a D3 peptide derived from human

XX high mol.wt. kininogen (HK) domain 3 (see AAY95426). The D3 peptide

XX inhibits endothelial cell proliferation and thus possesses

XX anti-angiogenic activity. It is an example of D3 peptides of the

XX invention (see AAY95405-26) that are analogues of certain sites in

XX the HK domain 3, in this case amino acid residues Leu331-Tyr338,

XX and in which native cysteine residues may be replaced by alanine

XX residues. The peptides inhibit endothelial cell proliferation and

XX may also induce endothelial cell apoptosis. Compositions including

XX the peptides are used in claimed methods for inhibiting angiogenesis,

XX inhibiting endothelial cell proliferation, and inducing endothelial

XX cell apoptosis. Cancer, rheumatoid arthritis, and ocular disorders

XX characterized by undesired vascularization of the retina are treated.

XX The IC50 value for the present peptide was 42 uM for inhibition of

XX fibroblast growth factor-induced HUVEC cell proliferation.

XX Sequence 16 AA;

XX Query Match

XX Best Local Similarity 100.0%; Score 18; DB 21; Length 16;

XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX QY 1 SLDA 4

XX 8 SLDA 11

AC AAR36604;

XX 02-SEP-1993 (first entry)

XX N-terminus of 37 kD acute phase protein.

XX Infection; tissue damage; assay kit; marker; bovine.

XX Bos taurus.

XX WO9309142-A.

XX 13-MAY-1993.

XX 06-NOV-1992; 92WO-CA00476.

XX 05-NOV-1992; 92US-0971559.

XX (UYSA-) UNIV SASKATCHEWAN.

XX Campos M, Redmond MJ, Harland R;

XX WPI; 1993-167625/20.

XX Assay kit for detection and prognosis of infection and tissue damage

XX in mammals - using acute phase proteins identified in bovine

XX subjects suffering from infection or induced tissue damage and their

XX specific antibodies

XX Claim 4; Page 23; 34pp; English.

XX The sequence shown is the N-terminal sequence of a 37 kD acute

XX phase protein isolated from bovine serum. The protein provides a

XX convenient marker for identifying the presence of infection prior to

XX the onset of observable symptoms, esp. in feedlot animals. Serious

XX illnesses may be thwarted and economic losses avoided. The protein

XX allows for monitoring of disease states and to identify contaminated

XX meat. X in the sequence defines an undetermined amino acid residue.

XX See also AAR36602-3.

XX Sequence 19 AA;

XX Query Match

XX Best Local Similarity 100.0%; Score 18; DB 14; Length 19;

XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX QY 1 SLDA 4

XX 5 SLDA 8

XX RESULT 36

XX AAY91397

XX ID AAY91397 standard; Protein; 19 AA.

XX AAY91397;

XX 29-JUN-2000 (first entry)

XX Human secreted protein sequence encoded by gene 23 SEQ ID NO:118.

XX Human; secreted protein; diagnosis; neuroprotective; nootropic;

XX neuroleptic; antianal; cerebroprotective; immunomodulatory;

XX anti-microbial; cardiact; cytostatic; antiinflammatory; haemostatic;

XX anticonvulsant; vasotropic; vaccine; gene therapy; anti-sense therapy;

XX lymphoma; demyelinating disease; autoimmunity; cancer; inflammation;

XX aneurysm; haemorrhage; Alzheimer's disease; Parkinson's disease;

XX Huntington's disease; Tourette syndrome; multiple sclerosis; meningitis;

XX ischaemia; mania; dementia; obsessive compulsive disorder;

XX viral prophylaxis; developmental disorder; sexually-linked disorder;

XX cardiovascular disorder; food additive; preservative.

OS Homo sapiens.  
 XX MO200011014-A1.  
 XX  
 XX  
 PD 02-MAR-2000.  
 XX  
 PF 24-AUG-1999; 99MO-US19330.  
 XX  
 PR 25-AUG-1998; 98US-0097917.  
 PR 31-AUG-1998; 98US-0098634.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Moore PA, Ruben SM, Olsen HS, Shi Y, Rosen CA, Florence KA;  
 PI Soppet DR, Lafleur DM, Endress GA, Edner R, Komatsoulis G;  
 PI Duan RD;  
 DR WPI: 2000-224656/19.  
 DR N-PSDB; AAA26332.  
 XX  
 PT Novel secreted proteins and corresponding DNA molecules that can be  
 PT used to prevent, treat and diagnose disease in humans, for example,  
 PT Alzheimer's, cancer, and immune disorders -  
 PS  
 PS Claim 11; Page 383; 416pp; English.  
 CC The polynucleotide sequences given in AAA26281 to AAA26336 encode the  
 CC human secreted proteins given in AAY91346 to AAY91449. The human secreted  
 CC proteins can have activities based on the tissues and cells they are  
 CC expressed in. Examples of the activities are: neuroprotective; neurotropic;  
 CC neuroleptic; anti-malignant; cerebroprotective; immunomodulatory;  
 CC anti-microbial; cardiant; cytostatic; anti-inflammatory; haemostatic;  
 CC anti-convasant; and vasotropic. The polynucleotides and proteins may be  
 CC used to prevent, treat or ameliorate a medical condition, e.g. by protein  
 CC or gene therapy. Conditions treatable by the proteins of the invention  
 CC include neural, reproductive, or immune disorders, especially  
 CC immunodeficiency, infection, lymphomas, demyelinating diseases,  
 CC auto-immunities, cancer, general microbial infection, inflammation,  
 CC aneurysms and haemorrhages. Specific examples include: Alzheimer's  
 CC disease; Parkinson's; Huntington's; Tourette syndrome; multiple  
 CC sclerosis; meningitis; ischaemia; prostate cancer; mania; dementia;  
 CC obsessive compulsive disorder and viral prophylaxis. The polynucleotides  
 CC and proteins can also be used in the detection of disorders associated  
 CC with the function of the protein, for example, the detection of  
 CC developmental disorders, sexually-linked disorders, or disorders of the  
 CC cardiovascular system. They may also be used as food additives or  
 CC preservatives. AAA26272 to AAA26280 and AAY91345 are sequences used in  
 CC the exemplification of the present invention.  
 CC  
 SQ Sequence 19 AA:  
 Query Match 100.0%; Score 18; DB 21; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 SLDA 4  
 Db 1 SLDA 4  
 XX  
 RESULT 37  
 AAY36691  
 ID AAY36691 standard; Protein: 20 AA.  
 XX  
 XX AAY36691;  
 AC  
 XX  
 DT 17-SEP-1999 (first entry)  
 XX  
 DE Fragment of human secreted protein encoded by gene 57.  
 XX  
 KW Human; secreted protein; cancer; tumour; developmental abnormality;  
 KW foetal deficiency; blood disorder; immune system disorder; inflammation;  
 KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;  
 XX

KW schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;  
 KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;  
 KW digestive disorder; endocrine disorder; infection; AIDS.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 PN MO9931117-A1.  
 XX  
 PD 24-JUN-1999.  
 XX  
 PF 17-DEC-1998; 98MO-US27059.  
 XX  
 PR 19-DEC-1997; 97US-0068369.  
 PR 18-DEC-1997; 97US-0068006.  
 PR 18-DEC-1997; 97US-0068007.  
 PR 18-DEC-1997; 97US-0068008.  
 PR 18-DEC-1997; 97US-0068053.  
 PR 18-DEC-1997; 97US-0068054.  
 PR 18-DEC-1997; 97US-0068057.  
 PR 18-DEC-1997; 97US-0068064.  
 PR 18-DEC-1997; 97US-0070923.  
 PR 19-DEC-1997; 97US-0068169.  
 PR 19-DEC-1997; 97US-0068365.  
 PR 19-DEC-1997; 97US-0068367.  
 PR 19-DEC-1997; 97US-0068368.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Carter KC, Duan RD, Feng P, Ferrie AM, Florence C;  
 PI Florence K, Greene JM, Janat F, Kyew H, Moore PA;  
 PI Ni J, Rosen CA, Ruben SM, Shi Y, Soppet DR, Wei Y;  
 PI Yu G;  
 DR WPI: 1999-418749/35.  
 DR  
 XX  
 PT  
 PT  
 PS  
 PS Disclosure; Page 522; 537pp; English.  
 CC AAX97916 to AAX98029 represent 110 isolated human secreted protein  
 CC genes. AAY36224 to AAY36727 represent the secreted proteins encoded by  
 CC the 110 human genes. The genes and their corresponding secreted  
 CC polypeptides are useful for preventing, treating or ameliorating medical  
 CC conditions, e.g. by protein or gene therapy. Also pathological conditions  
 CC can be diagnosed by determining the amount of the new polypeptides in a  
 CC sample or by determining the presence of mutations in the new genes.  
 CC Specific uses are described for each of the 110 genes, based on which  
 CC tissues they are most highly expressed in, and include developing  
 CC products for the diagnosis or treatment of cancer, tumours, developmental  
 CC abnormalities and foetal deficiencies, blood disorders, diseases of the  
 CC immune system, autoimmune diseases, inflammation, allergies, Alzheimer's  
 CC and cognitive disorders, schizophrenia, arthritis, asthma, psoriasis,  
 CC sepsis, skin disorders, atherosclerosis, diabetes, cardiovascular  
 CC disorders, kidney disorders, digestive/endocrine disorders, infections  
 CC and AIDS. The polypeptides are also useful for identifying their binding  
 CC partners. The sequences given in AAX97907 to AAX97915 and AAY36223 are  
 CC used in the exemplification of the present invention.  
 CC  
 SQ Sequence 20 AA:  
 Query Match 100.0%; Score 18; DB 20; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 SLDA 4  
 Db 14 SLDA 17  
 XX  
 RESULT 38  
 AAR33356  
 ID AAR33356 standard; peptide: 21 AA.  
 XX



XX cytostatic; antidiabetic; neuroprotective; cerebroprotective;  
 KW ophthalmological; insulin; receptor; gene therapy; diabetes;  
 KW insulin-like growth factor-1; IGF-1; tumour; prostate; breast;  
 KW diabetic retinopathy; neurological diseases; stroke;  
 KW diabetic neuropathy.  
 OS Synthetic.  
 XX  
 PN WO200172771-A2.  
 PD 04-OCT-2001.  
 XX  
 PF 29-MAR-2000; 2000MO-US08528.  
 XX  
 PR 29-MAR-2000; 2000MO-US08528.  
 XX  
 PA (DGIB-) DGI BIOTECHNOLOGIES LLC.  
 XX (NOVO) NOVO NORDISK AS.  
 PI Beasley J, Blume AJ, Schaeffer L, Pillutia R, Brandt J;  
 PI Brissette R, Spletzer J, Cheng W, Ostergaard S, Mandeckl WS;  
 PI Hansen PH, Raveira M, Hsiao K;  
 XX WPI: 2002-025774/03.  
 DR  
 XX  
 PT Modulating insulin activity in mammalian cells, for treating e.g.  
 PT diabetes and tumours, comprises using peptides that bind to insulin or  
 PT insulin-like growth factor receptors -  
 XX  
 XX Disclosure; Figure 1K-1; 390pp; English.  
 PS  
 XX  
 CC The invention relates to a method of modulating insulin activity in  
 CC mammalian cells by administering a peptide that binds the insulin  
 CC receptor (IR). A composition containing a peptide, optionally expressed  
 CC from gene therapy vectors, that binds to Site 1 of IR and an insulin  
 CC agonist are useful for treating diabetes. Also, peptides that are  
 CC antagonists of the insulin-like growth factor-1 (IGF-1) receptor are  
 CC useful for treating insulin-like growth factor (IGF)-sensitive tumours  
 CC (e.g. of prostate and breast) and diabetic retinopathy, while IGF-1  
 CC receptor agonists are useful for treating neurological diseases,  
 CC including stroke and diabetic neuropathy. The peptides are also useful in  
 CC screening for compounds that bind to IR or IGF-1 receptor, potential  
 CC therapeutics and research reagents. AAU8034-AAU90957 represent IR  
 CC and/or IGF-1 receptor-binding peptides and related amino acid sequences  
 CC of the invention.  
 CC  
 XX  
 SQ Sequence 26 AA:  
 XX  
 QY  
 Db 1 SLDA 4  
 1111  
 20 SLDA 23  
 23  
 RESULT 41  
 AAM10854  
 ID AAM10854 standard; peptide; 29 AA.  
 XX  
 AC AAM10854;  
 XX  
 DT 07-NOV-1997 (first entry)  
 XX  
 DE MAb anti-HBsAg binder sequence, FC11, from R26 library.  
 XX  
 KW Functional surrogate; analyte; affinity receptor; immunoreactive group;  
 KW mAb; homogenous immunoassay; detection; diagnostic analyte; Chlamydia;  
 KW cardiac marker; tumour marker; allergen; hormone; fertility; myoglobin;  
 KW pregnancy; infectious disease; ferritin; myosin light chain; tropomyosin;  
 KW follicle stimulating hormone; human; growth hormone; immunoglobulin E;

KW prolactin; parathyroid hormone; placental lactogen; hepatitis antigen;  
 KW antibody; chorionic gonadotropin; luteinising hormone; cytomegalovirus;  
 KW streptococcus; rubella; toxoplasma; DX-MB; prostate-specific antigen;  
 KW carcinoembryonic antigen; alpha-fetoprotein; herpes virus; CA125.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9641172-A1.  
 PD 19-DEC-1996.  
 XX  
 PF 07-JUN-1996; 96WO-US10498.  
 XX  
 PR 07-JUN-1995; 95US-0476375.  
 XX  
 PA (CYTO-) CYTOGEN CORP.  
 XX  
 PI Carter JM, Lee-Own FV;  
 PI  
 XX  
 DR WPI: 1997-077284/07.  
 XX  
 PT Labelled functional surrogate of an analyte - useful as competitor  
 PT molecule in affinity assays, esp. for detecting large macromolecules  
 PT such as ferritin  
 XX  
 XX  
 PS Claim 51; Page 95; 156pp; English.  
 XX  
 CC This sequence represents a monoclonal anti-hepatitis B antigen binder  
 CC sequence from the R26 library which may be used in the conjugate of the  
 CC invention. The novel labelled conjugate comprises at least one label  
 CC attached to a functional surrogate of an analyte of interest. The  
 CC surrogate is capable of competing effectively with the analyte for a  
 CC limiting amount of an affinity receptor for the analyte. The conjugate  
 CC exhibits an activity that is altered upon interaction with the affinity  
 CC receptor and this activity can be measured and related to the amount of  
 CC the analyte present in a sample. Functional surrogates such as this have  
 CC an immunoreactive group that allows the surrogate to compete effectively  
 CC and with the analyte for a limiting amount of its affinity receptor.  
 CC Functional surrogates are able to mimic naturally occurring analyses.  
 CC They can be labelled for use in standard competitive affinity assays  
 CC (esp. homogenous immunoassays) for detecting large macromolecules such  
 CC as polypeptides, polysaccharides, polynucleotides, glycoproteins and  
 CC lipid-containing macromolecules, as well as small haptens. Typical  
 CC diagnostic analyses for detection include cardiac or tumour markers,  
 CC allergens, hormones related to fertility-pregnancy or analyses associated  
 CC with infectious disease. In particular, the assays are useful for  
 CC detecting ferritin, follicle stimulating hormone, human growth hormone,  
 CC immunoglobulin E, prolactin, parathyroid hormone, human placental  
 CC lactogen, hepatitis antigens or antibodies against them, human  
 CC chorionic gonadotropin, human luteinising hormone, cytomegalovirus,  
 CC Chlamydia, Streptococcus a, rubella, toxoplasma, herpes virus, DX-MB,  
 CC myoglobin, myosin light chain, tropomyosin, carcinoembryonic antigen,  
 CC alpha-fetoprotein, prostate-specific antigen and CA125 (a tumour marker).  
 CC  
 XX  
 SQ Sequence 29 AA:  
 XX  
 QY  
 Db 1 SLDA 4  
 1111  
 10 SLDA 13  
 13  
 RESULT 42  
 AAR71050  
 ID AAR71050 standard; Peptide; 30 AA.  
 XX  
 AC AAR71050;  
 XX  
 DT 15-NOV-1995 (first entry)  
 XX

DE Draculin peptide #13.  
 XX  
 KW Anticoagulant; Draculin; saliva; vampire bat; inhibition; factor IX;  
 KW factor X; blood coagulation cascade; serine protease inhibitor; PMFS;  
 KW acute myocardial infarction; deep vein thrombosis; pulmonary embolism;  
 KW unstable angina; transient ischemic attack; peripheral vascular; DFP;  
 KW bypass occlusion; disseminated intravascular coagulation.  
 XX  
 OS Desmodus rotundus.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1  
 FT /note= "undefined amino acid"  
 XX  
 PN MO9505836-A.  
 XX  
 PD 02-MAR-1995.  
 XX  
 PF 22-AUG-1994; 94WO-US09488.  
 XX  
 PR 20-AUG-1993; 93US-0109807.  
 XX  
 PA (RHON ) RHONE POULENC RORER PHARM INC.  
 PI Aptiz-Castro R, Beguin S, Hemker HC, Holt JC, Lynch K;  
 PI Hemker H;  
 XX  
 DR WPI; 1995-106668/14.  
 XX  
 PT New anticoagulant protein, draculin, from vampire bat saliva -  
 PT and related nucleic acid, vectors, transformed cells and probes,  
 PT for treating myocardial infarction, etc  
 XX  
 PS Example 8; Page 34; 107pp; English.  
 XX  
 CC The sequences given in AAR71038-84 represent peptide fragments  
 CC derived from the anticoagulant protein, Draculin. Draculin  
 CC is isolated from the saliva of vampire bats and is distinct from  
 CC other known anticoagulants in that it inhibits both factors IX and X  
 CC of the blood coagulation cascade. The anticoagulant activity is not  
 CC inhibited by serine protease inhibitors such as PMFS or DFP. The  
 CC molecular weight of Draculin ranges from 75-90 kD and may be used in  
 CC treatment of acute myocardial infarction, deep vein thrombosis,  
 CC pulmonary embolism, unstable angina, transient ischemic attacks,  
 CC peripheral vascular or bypass occlusions and disseminated intravascular  
 CC coagulation. These peptide fragments were derived from Draculin by  
 CC cleavage with chymotrypsin and overlapping peptide fragments were  
 CC compared to determine the complete amino acid sequence of Draculin.  
 CC  
 SQ Sequence 30 AA;  
 Query Match 100.0%; Score 18; DB 16; Length 30;  
 Best Local Similarity 100.0%; Pred. No. 4.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 SLDA 4  
 | | | |  
 Db 6 SLDA 9  
 RESULT 43  
 AAW31683  
 ID AAW31683 standard; Protein; 30 AA.  
 XX  
 AC AAW31683;  
 XX  
 XX 27-MAR-1998 (first entry)  
 DE Clostridium perfringens type A enterotoxin (residues 291-319).  
 XX  
 KW Clostridium perfringens type A enterotoxin; toxoid; vaccine;  
 KW food poisoning.  
 XX

OS Clostridium perfringens.  
 OS Synthetic.  
 XX  
 PN US5695956-A.  
 XX  
 PD 09-DEC-1997.  
 XX  
 PF 15-MAR-1994; 94US-0213452.  
 XX  
 PR 26-NOV-1990; 90US-0618541.  
 PR 22-APR-1992; 92US-0874982.  
 PR 15-MAR-1994; 94US-0213452.  
 XX  
 PA (UYPI-) UNIV PITTSBURGH.  
 PI Hanna PC, McLane BA, Mletzner TN;  
 PI WPI; 1998-041298/04.  
 DR  
 XX  
 PF DNA molecule encoding Clostridium perfringens enterotoxin fragment -  
 PF useful as toxoid in vaccine for preventing food poisoning  
 XX  
 PS Claim 14; Columns 19-20; 24pp; English.  
 XX  
 CC This sequence corresponds to a carboxy-terminal protein fragment  
 CC (amino acid residues 290-319) of the Clostridium perfringens type A  
 CC enterotoxin protein. Plasmids containing E. coli expression vectors and  
 CC Clostridium perfringens type A enterotoxin gene fragment encoding amino  
 CC acid residues 171-319 or 290-319 can be used to produce Clostridium  
 CC perfringens type A enterotoxin toxoids. A synthetic peptide corresponding  
 CC to these residues 290-319 can also be prepared to be used as a toxoid.  
 CC These toxoids irreversibly bind to and saturate receptor sites on  
 CC intestinal membranes, thus effectively competing for these receptor sites  
 CC with Clostridium perfringens type A enterotoxin. The synthetic peptide  
 CC and the recombinant carboxy-terminal fragment, are useful as toxoids in  
 CC a vaccine for preventing C. perfringens food poisoning, or in a  
 CC composition to treat such food poisoning.  
 CC  
 SQ Sequence 30 AA;  
 Query Match 100.0%; Score 18; DB 19; Length 30;  
 Best Local Similarity 100.0%; Pred. No. 4.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 SLDA 4  
 | | | |  
 Db 1 SLDA 4  
 RESULT 44  
 AAY95418  
 ID AAY95418 standard; Peptide; 32 AA.  
 XX  
 AC AAY95418;  
 XX  
 XX 25-SEP-2000 (first entry)  
 DE Anti-angiogenic D3 peptide.  
 XX  
 KW Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;  
 KW endothelial cell proliferation; apoptosis; cancer; ocular disorder;  
 KW rheumatoid arthritis; cyostatic; antiarthritic; antiinematic;  
 KW therapy; human; D3 peptide; cyclic.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200035407-A2.  
 XX  
 PD 22-JUN-2000.  
 XX  
 PF 02-DEC-1999; 99WO-US28465.  
 XX  
 PR 16-DEC-1998; 98US-0112427.  
 XX



```

XX (UTEM ) UNIV TEMPLE.
PA (MCCR/) MCCRAE R K.
XX
PI MCCRAE RK;
XX
DR WPI: 2000-442247/38.
XX
PT Composition for inhibiting angiogenesis and endothelial cell
PT proliferation, inducing endothelial cell apoptosis and treating cancer,
PT rheumatoid arthritis, and ocular disorders comprises a kininogen domain
PT 3 analog
XX
PS Claim 18; Page 29; 44pp; English.
XX
CC The present sequence is that of a D3 peptide derived from human
CC high mol.wt. kininogen (HK) domain 3 (see AA95426). The D3 peptide
CC inhibits endothelial cell proliferation and thus possesses
CC anti-angiogenic activity. It is an example of D3 peptides of the
CC invention (see AA95405-26) that are analogues of certain sites in
CC the HK domain 3. In this case amino acid residues Leu331-Tyr338,
CC where native cysteine residues may be replaced by alanine residues.
CC The peptides inhibit endothelial cell proliferation and may also
CC induce endothelial cell apoptosis. Compositions including the
CC peptides are used in claimed methods for inhibiting angiogenesis,
CC inhibiting endothelial cell proliferation, and inducing endothelial
CC cell apoptosis. Cancer, rheumatoid arthritis, and ocular disorders
CC characterized by undesired vascularization of the retina are treated.
XX
SQ Sequence 32 AA:
XX
Query Match 100.0%; Score 18; DB 21; Length 32;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SLDA 4
DB 12 SLDA 15
XX
RESULT 45
AAW10869
ID AAW10869 standard; peptide; 33 AA.
XX
AC AAW10869;
XX
DT 07-NOV-1997 (first entry)
XX
DE MAb anti-HBsAg binder sequence, FC11, from R26 library.
XX
KW Functional surrogate; analyte; affinity receptor; immunoreactive group;
KW mimic; homogeneous immunoassay; detection; diagnostic analyte; Chlamydia;
KW cardiac marker; tumour marker; allergen; hormone; fertility; myoglobin;
KW pregnancy; infectious disease; ferritin; myosin light chain; tropoin;
KW follicle stimulating hormone; human; growth hormone; immunoglobulin E;
KW prolactin; parathyroid hormone; placental lactogen; hepatitis antigen;
KW antibody; chorionic gonadotropin; luteinising hormone; cytomegalovirus;
KW Streptococcus; rubella; toxoplasma; DK-WB; prostate-specific antigen;
KW carcinoembryonic antigen; alpha-fetoprotein; herpes virus; CA125.
XX
OS Synthetic.
XX
PN W09641172-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96MO-US10498.
XX
PR 07-JUN-1995; 95US-0476375.
XX
PA (CYTO-) CYTOGEN CORP.
XX
PI Carter JM, Lee-Own FV;

```

```

XX WPI: 1997-077284/07.
DR N-PSDB; AAF48665.
XX
PT Labelled functional surrogate of an analyte - useful as competitor
PT molecule in affinity assays, esp. for detecting large macromolecules
PT such as ferritin
XX
PS Claim 51; Page 100; 156pp; English.
XX
CC This sequence represents a monoclonal anti-hepatitis B antigen binder
CC sequence from the R26 library which may be used in the conjugate of the
CC invention. The novel labelled conjugate comprises at least one label
CC attached to a functional surrogate of an analyte of interest. The
CC surrogate is capable of competing effectively with the analyte for a
CC limiting amount of an affinity receptor for the analyte. The conjugate
CC exhibits an activity that is altered upon interaction with the affinity
CC receptor and this activity can be measured and related to the amount of
CC the analyte present in a sample. Functional surrogates such as this have
CC an immunoreactive group that allows the surrogate to compete effectively
CC and with the analyte for a limiting amount of its affinity receptor.
CC Functional surrogates are able to mimic naturally occurring analytes.
CC They can be labelled for use in standard competitive affinity assays
CC (esp. homogeneous immunoassays) for detecting large macromolecules such
CC as polypeptides, polysaccharides, polynucleotides, glycoproteins and
CC lipid-containing macromolecules, as well as small haptens. Typical
CC diagnostic analytes for detection include cardiac or tumour markers,
CC allergens, hormones related to fertility/pregnancy or analytes associated
CC with infectious disease. In particular, the assays are useful for
CC detecting ferritin, follicle stimulating hormone, human growth hormone,
CC immunoglobulin E, prolactin, parathyroid hormone, human placental
CC lactogen, hepatitis antigens or antibodies against them, human
CC chorionic gonadotropin, human luteinising hormone, cytomegalovirus,
CC Chlamydia, Streptococcus a, rubella, toxoplasma, herpes virus, DK-WB,
CC myoglobin, myosin light chain, tropoin, carcinoembryonic antigen,
CC alpha-fetoprotein, prostate-specific antigen and CA125 (a tumour marker).
XX
SQ Sequence 33 AA:
XX
Query Match 100.0%; Score 18; DB 18; Length 33;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SLDA 4
DB 14 SLDA 17

```

Search completed: February 6, 2003, 11:19:40  
 Job time : 29.8333 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: February 6, 2003, 11:16:04 : Search time 9.33333 Seconds  
(without alignments)  
12.610 Million cell updates/sec

Title: PAT943-4  
Perfect score: 18  
Sequence: 1 slida 4

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues  
Total number of hits satisfying chosen parameters: 597

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_AA:  
1: /cgn2.6/ptodata/1/1aa/5A\_COMB.pep.\*  
2: /cgn2.6/ptodata/1/1aa/5B\_COMB.pep.\*  
3: /cgn2.6/ptodata/1/1aa/5A\_COMB.pep.\*  
4: /cgn2.6/ptodata/1/1aa/5B\_COMB.pep.\*  
5: /cgn2.6/ptodata/1/1aa/PCTUS\_COMB.pep.\*  
6: /cgn2.6/ptodata/1/1aa/Dackfilest1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	100.0	7	1	US-08-136-743B-48
2	18	100.0	7	3	US-09-040-216-20
3	18	100.0	7	4	US-09-001-984C-53
4	18	100.0	7	4	US-09-426-568A-10
5	18	100.0	8	1	US-08-526-710-12
6	18	100.0	8	3	US-08-862-855-12
7	18	100.0	8	4	US-09-226-906-12
8	18	100.0	8	4	US-09-227-906-12
9	18	100.0	13	1	US-08-464-531-96
10	18	100.0	13	2	US-08-461-598-96
11	18	100.0	13	3	US-08-322-137-96
12	18	100.0	13	3	US-08-582-333A-32
13	18	100.0	13	1	US-08-213-452-3
14	18	100.0	63	4	US-08-887-534A-94
15	18	100.0	64	2	US-08-209-521-19
16	18	100.0	64	2	US-08-209-521-20
17	18	100.0	64	4	US-08-961-810-129
18	18	100.0	64	4	US-08-961-810-130
19	18	100.0	64	4	US-08-352-902D-129
20	18	100.0	64	4	US-08-352-902D-130
21	18	100.0	73	4	US-08-469-260A-33
22	18	100.0	84	1	US-08-798-897-16
23	18	100.0	84	2	US-08-978-523-16
24	18	100.0	85	4	US-09-134-001C-3985
25	18	100.0	90	4	US-09-186-276B-59
26	18	100.0	90	4	US-08-842-445-59
27	18	100.0	90	4	US-09-186-188B-59

28	18	100.0	105	4	US-08-936-165A-515	Sequence 515, App
29	18	100.0	109	4	US-09-099-041A-6	Sequence 6, Appl
30	18	100.0	109	4	US-09-199-637A-231	Sequence 231, App
31	18	100.0	109	4	US-09-245-281-6	Sequence 6, Appl
32	18	100.0	110	4	US-09-207-359B-6	Sequence 6, Appl
33	18	100.0	114	4	US-09-437-054A-2	Sequence 2, Appl
34	18	100.0	121	2	US-08-470-670A-15	Sequence 15, Appl
35	18	100.0	121	4	US-08-461-511A-15	Sequence 15, Appl
36	18	100.0	124	4	US-09-134-001C-2929	Sequence 2929, Ap
37	18	100.0	149	1	US-08-213-452-2	Sequence 2, Appl
38	18	100.0	152	2	US-08-606-143-18	Sequence 18, Appl
39	18	100.0	166	4	US-08-954-395A-16	Sequence 16, Appl
40	18	100.0	167	4	US-09-069-023-6	Sequence 6, Appl
41	18	100.0	170	1	US-08-081-448-8	Sequence 8, Appl
42	18	100.0	170	2	US-08-470-670A-9	Sequence 9, Appl
43	18	100.0	170	4	US-08-461-511A-9	Sequence 9, Appl
44	18	100.0	170	4	US-09-271-014A-8	Sequence 8, Appl
45	18	100.0	170	5	PCT-US94-07089-9	Sequence 9, Appl

## ALIGNMENTS

RESULT 1  
US-08-136-743B-48  
Sequence 48, Application US/08136743B  
Patent No. 5459063  
GENERAL INFORMATION:  
APPLICANT: Barry S. Cooperman, Harvey Rubin,  
APPLICANT: Jerome Salem, and Allison L. Fisher  
TITLE OF INVENTION: "Plasmodium falciparum Ribonu-  
TITLE OF INVENTION: cleotide Reductase, DNA Sequences Therefor and Peptide inh  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: The University of Pennsylvania  
STREET: Suite 330  
STREET: 3700 Market Street  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: U.S.A.  
ZIP: 19104-3246  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/136,743B  
FILING DATE: 10/14/93  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Monaco, Daniel A.  
REGISTRATION NUMBER: 30,480  
REFERENCE/DOCKET NUMBER: 3957-10  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-8383  
TELEFAX: (215) 568-5549  
TELEX: No. 5459063e  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-136-743B-48  
Query Match 100.0%; Score 18; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0;  
Gaps 0;  
OY 1 SLDA 4  
DB 2 SLDA 5

```

RESULT 2
US-09-040-216-20
; Sequence 20, Application US/09040216
; Patent No. 6030942
; GENERAL INFORMATION:
; APPLICANT: COOPERMAN, ET AL., BARRY
; TITLE OF INVENTION: PEPTIDES, PEPTIDE ANALOGS, PEPTIDOMIMETICS, AND OTHER
; TITLE OF INVENTION: SMALL MOLECULES USEFUL FOR INHIBITING THE ACTIVITY OF
; FILE REFERENCE: 9596-63U1
; CURRENT APPLICATION NUMBER: US/09/040,216
; PRIOR FILING DATE: 1998-03-17
; EARLIER APPLICATION NUMBER: 08/919,748
; EARLIER FILING DATE: 1997-08-28
; EARLIER APPLICATION NUMBER: 60/025,146
; EARLIER FILING DATE: 1996-08-30
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 20
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; FEATURE: Ribonucleotide reductase inhibitor peptide
; OTHER INFORMATION: residue 1: MOD_RES: ACETYLATION
US-09-040-216-20

```

```

Query Match
Best Local Similarity 100.0%; Score 18; DB 3; Length 7;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 SIDA 4
    |||
Db 2 SIDA 5

```

```

RESULT 3
US-09-001-984C-53
; Sequence 53, Application US/09001984C
; Patent No. 6245331
; GENERAL INFORMATION:
; APPLICANT: Laal, Suman
; APPLICANT: Zolla-Pazner, Susan
; TITLE OF INVENTION: EARLY DETECTION OF MYCOBACTERIAL DISEASE
; FILE REFERENCE: NYU-011
; CURRENT APPLICATION NUMBER: US/09/001,984C
; PRIOR FILING DATE: 1997-12-31
; PRIOR APPLICATION NUMBER: 60/034,003
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 53
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis strain H37Rv
US-09-001-984C-53

```

```

Query Match
Best Local Similarity 100.0%; Score 18; DB 4; Length 7;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 SIDA 4
    |||
Db 4 SIDA 7

```

```

RESULT 4
US-09-426-568A-10

```

```

; Sequence 10, Application US/09426568A
; Patent No. 6348643
; GENERAL INFORMATION:
; APPLICANT: Kakefuda, Genichi
; APPLICANT: Costello, Colleen
; APPLICANT: Sun, Ming
; APPLICANT: Hu, Weiming
; TITLE OF INVENTION: Genes and Vectors for Confering Herbicide Resistance
; TITLE OF INVENTION: In Plants
; FILE REFERENCE: 008103/195497
; CURRENT APPLICATION NUMBER: US/09/426,568A
; PRIOR FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/106,239
; PRIOR FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 10
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Arabidopsis sp.
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(7)
; OTHER INFORMATION: N-terminal sequence of AHAS small subunit peptide
; NAME/KEY: SITE
; LOCATION: (1)..(2)
; OTHER INFORMATION: Thrombin cleavage site
US-09-426-568A-10

```

```

Query Match
Best Local Similarity 100.0%; Score 18; DB 4; Length 7;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 SIDA 4
    |||
Db 2 SIDA 5

```

```

RESULT 5
US-08-526-710-12
; Sequence 12, Application US/08526710
; Patent No. 5622699
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Method of Identifying Molecules That
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESS: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/526,710
; FILING DATE: 11-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1779
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 12:

```

SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-526-710-12

Query Match 100.0%; Score 18; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
|||  
DB 4 SLDA 7

RESULT 6  
US-08-862-855-12  
Sequence 12, Application US/08862855  
Patent No. 6068829  
GENERAL INFORMATION:  
APPLICANT: Ruoslahti, Erkki  
APPLICANT: Pasqualini, Renata  
TITLE OF INVENTION: Method of Identifying Molecules That  
TITLE OF INVENTION: Home to a Selected Organ In Vivo  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell & Flores LLP  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: United States  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/862,855  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION NUMBER: US 08/526,710  
FILING DATE: 11-SEP-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/813,273  
FILING DATE: 10-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 2621  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-862-855-12

Query Match 100.0%; Score 18; DB 3; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
|||  
DB 4 SLDA 7

RESULT 7

US-09-226-985-12  
Sequence 12, Application US/09226985  
Patent No. 6296832  
GENERAL INFORMATION:  
APPLICANT: Ruoslahti, Erkki  
APPLICANT: Pasqualini, Renata  
TITLE OF INVENTION: Molecules That Home to a Selected Organ In Vivo  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell & Flores LLP  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: United States  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/226,985  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/526,710  
FILING DATE: 11-SEP-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/813,273  
FILING DATE: 10-MAR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/862,855  
FILING DATE: 23-MAY-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 3423  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-226-985-12

Query Match 100.0%; Score 18; DB 4; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.9e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
|||  
DB 4 SLDA 7

RESULT 8  
US-09-227-906-12  
Sequence 12, Application US/09227906  
Patent No. 6306365  
GENERAL INFORMATION:  
APPLICANT: Ruoslahti, Erkki  
APPLICANT: Pasqualini, Renata  
TITLE OF INVENTION: Method of Identifying Molecules That  
TITLE OF INVENTION: Home to a Selected Organ In Vivo  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell & Flores LLP  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: United States

ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/227,906  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/526,710  
FILING DATE: 11-SEP-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/813,273  
FILING DATE: 10-MAR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/862,855  
FILING DATE: 23-MAY-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 3424  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-227-906-12

Query Match  
Best Local Similarity 100.0%; Score 18; DB 4; Length 8;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
DB 4 SLDA 7

RESULT 9  
US-08-464-531-96  
Sequence 96, Application US/08464531  
Patent No. 5789184  
GENERAL INFORMATION:  
APPLICANT: FOLMKS, Dana M.  
APPLICANT: BROACH, Jim  
APPLICANT: MANFREDI, John  
APPLICANT: KLEIN, Christine  
APPLICANT: MURPHY, Andrew J.  
APPLICANT: PAUL, Jeremy  
APPLICANT: TRUEHEART, Joshua  
TITLE OF INVENTION: YEAST CELLS ENGINEERED TO PRODUCE  
NUMBER OF SEQUENCES: 119  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEIMARK  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,531  
FILING DATE: 05-JUN-1995

CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/322,137  
FILING DATE: 13-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/309,313  
FILING DATE: 20-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/190,328  
FILING DATE: 31-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/041,431  
FILING DATE: 31-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: COOPER, Iver P.  
REGISTRATION NUMBER: 28,005  
REFERENCE/DOCKET NUMBER: FOLMKS-2G  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 96:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-464-531-96

Query Match  
Best Local Similarity 100.0%; Score 18; DB 1; Length 13;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
DB 5 SLDA 8

RESULT 10  
US-08-461-598-96  
Sequence 96, Application US/08461598  
Patent No. 5876951  
GENERAL INFORMATION:  
APPLICANT: FOLMKS, Dana M.  
APPLICANT: BROACH, Jim  
APPLICANT: MANFREDI, John  
APPLICANT: KLEIN, Christine  
APPLICANT: MURPHY, Andrew J.  
APPLICANT: PAUL, Jeremy  
APPLICANT: TRUEHEART, Joshua  
TITLE OF INVENTION: YEAST CELLS ENGINEERED TO PRODUCE  
NUMBER OF SEQUENCES: 119  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEIMARK  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/461,598  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/322,137  
FILING DATE: 13-OCT-1994  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/309,313  
FILING DATE: 20-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/190,328  
FILING DATE: 31-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/041,431  
FILING DATE: 31-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: COOPER, Iver P.  
REGISTRATION NUMBER: 28,005  
REFERENCE/DOCKET NUMBER: FOLMKES-2F  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 96:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-461-598-96

Query Match 100.0%; Score 18; DB 2; Length 13;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
||||  
DB 5 SLDA 8

RESULT 11  
US-08-322-137-96  
Sequence 96, Application US/08322137  
Patent No. 6100042  
GENERAL INFORMATION:  
APPLICANT: FOLMKES, Dana M.  
APPLICANT: BROACH, Jim  
APPLICANT: MANFREDI, John  
APPLICANT: KLEIN, Christine  
APPLICANT: MURPHY, Andrew J.  
APPLICANT: PAUL, Jeremy  
APPLICANT: TRUEHEART, Joshua  
TITLE OF INVENTION: YEAST CELLS ENGINEERED TO PRODUCE  
TITLE OF INVENTION: PHEROMONE SYSTEM PROTEIN SURROGATES, AND USES THEREFOR  
NUMBER OF SEQUENCES: 119  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEIMARK  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/322,137  
FILING DATE: 13-OCT-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/309,313  
FILING DATE: 20-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/190,328  
FILING DATE: 31-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/041,431  
FILING DATE: 31-MAR-1993

ATTORNEY/AGENT INFORMATION:  
NAME: COOPER, Iver P.  
REGISTRATION NUMBER: 28,005  
REFERENCE/DOCKET NUMBER: FOLMKES-2C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 96:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-322-137-96

Query Match 100.0%; Score 18; DB 3; Length 13;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
||||  
DB 5 SLDA 8

RESULT 12  
US-08-582-333A-32  
Sequence 32, Application US/08582333A  
Patent No. 6255059  
GENERAL INFORMATION:  
APPLICANT: Klein, Christine A.  
APPLICANT: Murphy, Andrew J. M.  
TITLE OF INVENTION: Methods and Compositions for  
TITLE OF INVENTION: Identifying Receptor Effectors  
NUMBER OF SEQUENCES: 98  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Ascii(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/582,333A  
FILING DATE: 17-JAN-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Catherine J. Kara  
REGISTRATION NUMBER: 41,106  
REFERENCE/DOCKET NUMBER: CPI-012CPS  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-4214  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-582-333A-32

Query Match 100.0%; Score 18; DB 4; Length 13;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
||||  
DB 5 SLDA 8

RESULT 13  
US-08-213-452-3  
Sequence 3, Application US/08213452  
Patent No. 565956  
GENERAL INFORMATION:  
APPLICANT: McCrane, Bruce A.  
APPLICANT: Hanna, Philip C.  
APPLICANT: Metzner, Timothy A.  
TITLE OF INVENTION: Clostridium perfringens Type A  
TITLE OF INVENTION: Enterotoxin Toxoid and Methods of Preparation and Use As A  
TITLE OF INVENTION: Vaccine and Therapeutic Agent  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Lewis F. Gould, Jr.  
STREET: 1700 Market St. Suite 3232  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: US  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/213,452  
FILING DATE:  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Gould, Jr., Lewis F.  
REGISTRATION NUMBER: 25,057  
REFERENCE/DOCKET NUMBER: 107291-3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 575-6015  
TELEFAX: (215) 575-6000  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: C-terminal  
ORIGINAL SOURCE:  
ORGANISM: Clostridium perfringens  
STRAIN: NCTC 8239  
US-08-213-452-3

Query Match 100.0%; Score 18; DB 1; Length 30;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
DB 1 SLDA 4

RESULT 14  
US-08-887-534A-94  
Sequence 94, Application US/0887534A  
Patent No. 6455323  
GENERAL INFORMATION:  
APPLICANT: Hoiden, David W.  
TITLE OF INVENTION: ANTI-BACTERIAL METHODS AND MATERIALS  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive/6300 Sears Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America

ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/887,534A  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Rin-Laures, Li-Hsien  
REGISTRATION NUMBER: 33,547  
REFERENCE/DOCKET NUMBER: 28341/33996  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX: (312) 474-6600  
INFORMATION FOR SEQ ID NO: 94:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 63 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-887-534A-94

Query Match 100.0%; Score 18; DB 4; Length 63;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
DB 31 SLDA 34

RESULT 15  
US-08-209-521-19  
Sequence 19, Application US/08209521  
Patent No. 5922855  
GENERAL INFORMATION:  
APPLICANT: Liskay, Robert M.  
APPLICANT: Bronner, C. Eric  
APPLICANT: Baker, Sean M.  
APPLICANT: Bollag, Roni J.  
APPLICANT: Kolodner, Richard D.  
TITLE OF INVENTION: MAMMALIAN DNA MISMATCH REPAIR GENES  
TITLE OF INVENTION: hMLH1 AND hPMS1  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Kolisch, Hartwell, Dickinson, McCormack &  
STREET: 520 S.W. Yamhill, Suite 200  
CITY: Portland  
STATE: Oregon  
COUNTRY: US  
ZIP: 97204  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/209,521  
FILING DATE: 08-MAR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Van Rysseberghe, Pierre C.  
REGISTRATION NUMBER: 33,557  
REFERENCE/DOCKET NUMBER: OHSU 306A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (503) 224-6655  
TELEFAX: (503) 295-6679  
INFORMATION FOR SEQ ID NO: 19:



SEQUENCE CHARACTERISTICS:  
LENGTH: 64 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-209-521-19

Query Match 100.0%; Score 18; DB 2; Length 64;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
||||  
Db 24 SLDA 27

RESULT 16  
US-08-209-521-20  
Sequence 20, Application US/08209521  
Patent No. 5922855

GENERAL INFORMATION:  
APPLICANT: Liskay, Robert M.  
APPLICANT: Bronner, C. Eric  
APPLICANT: Baker, Sean M.  
APPLICANT: Bollag, Roni J.  
APPLICANT: Kolodner, Richard D.  
TITLE OF INVENTION: MAMMALIAN DNA MISMATCH REPAIR GENES  
TITLE OF INVENTION: hMLH1 AND hPMS1  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kolisch, Hartwell, Dickinson, McCormack &  
STREET: 520 S.W. Yamhill, Suite 200  
CITY: Portland  
STATE: Oregon  
COUNTRY: US  
ZIP: 97204

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/209,521  
FILING DATE: 08-MAR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Van Rysseberghe, Pierre C.  
REGISTRATION NUMBER: 33,557  
REFERENCE/DOCKET NUMBER: OHSU 306A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (503) 224-6655  
TELEFAX: (503) 295-6679  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 64 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-209-521-20

Query Match 100.0%; Score 18; DB 2; Length 64;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
||||  
Db 24 SLDA 27

RESULT 17

US-08-961-810-129  
Sequence 129, Application US/08961810  
Patent No. 6165713

GENERAL INFORMATION:  
APPLICANT: Liskay, Robert M.  
APPLICANT: Bronner, C. Eric  
APPLICANT: Baker, Sean M.  
APPLICANT: Bollag, Roni J.  
APPLICANT: Kolodner, Richard D.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS RELATING TO DNA  
TITLE OF INVENTION: MISMATCH REPAIR GENES  
NUMBER OF SEQUENCES: 134  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kolisch, Hartwell, Dickinson, McCormack &  
STREET: 520 S.W. Yamhill Street, Suite 200  
CITY: Portland  
STATE: Oregon  
COUNTRY: U.S.A.  
ZIP: 97204

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/961,810  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Van Rysseberghe, Pierre C.  
REGISTRATION NUMBER: 33,557  
REFERENCE/DOCKET NUMBER: OHSU 306B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (503) 224-6655  
TELEFAX: (503) 295-6679  
TELEX: 360619

INFORMATION FOR SEQ ID NO: 129:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 64 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-961-810-129

Query Match 100.0%; Score 18; DB 4; Length 64;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
||||  
Db 24 SLDA 27

RESULT 18  
US-08-961-810-130  
Sequence 130, Application US/08961810  
Patent No. 6165713

GENERAL INFORMATION:  
APPLICANT: Liskay, Robert M.  
APPLICANT: Bronner, C. Eric  
APPLICANT: Baker, Sean M.  
APPLICANT: Bollag, Roni J.  
APPLICANT: Kolodner, Richard D.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS RELATING TO DNA  
TITLE OF INVENTION: MISMATCH REPAIR GENES  
NUMBER OF SEQUENCES: 134  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kolisch, Hartwell, Dickinson, McCormack &  
STREET: 520 S.W. Yamhill Street, Suite 200  
CITY: Portland

Query Match 100.0%; Score 18; DB 4; Length 64;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

STATE: Oregon  
COUNTRY: U.S.A.  
ZIP: 97204  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/961,810  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Van Rysselberghe, Pierre C.  
REGISTRATION NUMBER: 33,557  
REFERENCE/DOCKET NUMBER: OHSU 306B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (503) 224-6655  
TELEFAX: (503) 295-6679  
TELEX: 360619  
INFORMATION FOR SEQ ID NO: 130:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 64 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-961-810-130

Query Match  
Best Local Similarity 100.0%; Score 18; DB 4; Length 64;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|||||  
DB 24 SLDA 27

RESULT 19  
US-08-352-902D-129  
Sequence 129, Application US/08352902D  
Patent No. 6191268  
GENERAL INFORMATION:  
APPLICANT: Liskay, Robert M.  
Bromner, C. Eric  
Baker, Sean M.  
Bollag, Roni J.  
Kolodner, Richard D.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS RELATING TO DNA  
MISMATCH REPAIR GENES  
NUMBER OF SEQUENCES: 149  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kolisch, Hartwell, Dickinson, McCormack & Heuser  
STREET: 520 S.W. Yamhill Street, Suite 200  
CITY: Portland  
STATE: Oregon  
COUNTRY: U.S.A.  
ZIP: 97204  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/352,902D  
FILING DATE: 09-Dec-1994  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Van Rysselberghe, Pierre C.  
REGISTRATION NUMBER: 33,557  
REFERENCE/DOCKET NUMBER: OHSU 306B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (503) 224-6655  
TELEFAX: (503) 295-6679  
TELEX: 360619  
INFORMATION FOR SEQ ID NO: 130:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 64 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 130:  
US-08-352-902D-130

Query Match  
Best Local Similarity 100.0%; Score 18; DB 4; Length 64;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

TELEPHONE: (503) 224-6655.  
TELEFAX: (503) 295-6679  
TELEX: 360619  
INFORMATION FOR SEQ ID NO: 129:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 64 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 129:  
US-08-352-902D-129

Query Match  
Best Local Similarity 100.0%; Score 18; DB 4; Length 64;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|||||  
DB 24 SLDA 27

RESULT 20  
US-08-352-902D-130  
Sequence 130, Application US/08352902D  
Patent No. 6191268  
GENERAL INFORMATION:  
APPLICANT: Liskay, Robert M.  
Bromner, C. Eric  
Baker, Sean M.  
Bollag, Roni J.  
Kolodner, Richard D.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS RELATING TO DNA  
MISMATCH REPAIR GENES  
NUMBER OF SEQUENCES: 149  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kolisch, Hartwell, Dickinson, McCormack & Heuser  
STREET: 520 S.W. Yamhill Street, Suite 200  
CITY: Portland  
STATE: Oregon  
COUNTRY: U.S.A.  
ZIP: 97204  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/352,902D  
FILING DATE: 09-Dec-1994  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Van Rysselberghe, Pierre C.  
REGISTRATION NUMBER: 33,557  
REFERENCE/DOCKET NUMBER: OHSU 306B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (503) 224-6655  
TELEFAX: (503) 295-6679  
TELEX: 360619  
INFORMATION FOR SEQ ID NO: 130:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 64 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 130:  
US-08-352-902D-130

Query Match  
Best Local Similarity 100.0%; Score 18; DB 4; Length 64;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
||||  
Db 24 SLDA 27

## RESULT 21

US-08-469-260A-33  
; Sequence 33, Application US/08469260A  
; Patent No. 6451578  
; GENERAL INFORMATION:  
; APPLICANT: JOHN N. SIMONS  
; APPLICANT: TAMI J. PILOT-MATIAS  
; APPLICANT: GEORGE J. DAMSON  
; APPLICANT: GEORGE G. SCHLAUDER  
; APPLICANT: SURESH M. DESAI  
; APPLICANT: THOMAS P. LEARY  
; APPLICANT: ANTHONY SCOTT MUEHRHOFF  
; APPLICANT: JAMES C. ERKER  
; APPLICANT: SHERI L. BULIK  
; APPLICANT: ISA K. MUSAHMAR  
; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS  
; TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE  
; NUMBER OF SEQUENCES: 716  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D  
; STREET: 100 ABBOTT PARK ROAD  
; CITY: ABBOTT PARK  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60064-3500  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/469,260A  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/424,550  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: POROMBSKI, PRISCILLA E.  
; REGISTRATION NUMBER: 33,207  
; REFERENCE/DOCKET NUMBER: 5527.PC.01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 708-937-6365  
; TELEFAX: 708-938-2623  
; INFORMATION FOR SEQ ID NO: 33:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 73 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-469-260A-33  
Query Match 100.0%; Score 18; DB 4; Length 73;  
Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
||||  
Db 58 SLDA 61

## RESULT 22

US-08-798-897-16  
; Sequence 16, Application US/08798897  
; Patent No. 5789201  
; GENERAL INFORMATION:

; APPLICANT: Guastella, John  
; TITLE OF INVENTION: Genes Coding For Bcl-Y, a Bcl-2  
; TITLE OF INVENTION: Homologue  
; NUMBER OF SEQUENCES: 53  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 New York Avenue, N.W., Suite 600  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/798,897  
; FILING DATE: February 11, 1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Esmond, Robert W.  
; REGISTRATION NUMBER: 32,893  
; REFERENCE/DOCKET NUMBER: 1483.0140001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-2600  
; TELEFAX: 202-371-2540  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 84 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-798-897-16  
Query Match 100.0%; Score 18; DB 1; Length 84;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
||||  
Db 70 SLDA 73

## RESULT 23

US-08-978-523-16  
; Sequence 16, Application US/08978523  
; Patent No. 5883229  
; GENERAL INFORMATION:  
; APPLICANT: Guastella, John  
; TITLE OF INVENTION: Genes Coding For Bcl-Y, a Bcl-2  
; TITLE OF INVENTION: Homologue  
; NUMBER OF SEQUENCES: 53  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 New York Avenue, N.W., Suite 600  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/978,523  
; FILING DATE: herewith  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/798,897  
; FILING DATE: February 11, 1997

```

; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1483.0140002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-978-523-16

Query Match
Best Local Similarity 100.0%; Score 18; DB 2; Length 84;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SIDA 4
   |||
Db 70 SIDA 73

RESULT 24
US-09-134-001C-3985
; Sequence 3985, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3985
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3985

Query Match
Best Local Similarity 100.0%; Score 18; DB 4; Length 85;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SIDA 4
   |||
Db 26 SIDA 29

RESULT 25
US-09-186-276B-59
; Sequence 59, Application US/09186276B
; Patent No. 6388173
; GENERAL INFORMATION:
; APPLICANT: Benfey, Philip
; APPLICANT: DiLaurenzio, Laura
; APPLICANT: Wyszoka-Diller, Joanna
; APPLICANT: Malsamy, Jocelyn E.
; APPLICANT: Pysn, Leonard
; APPLICANT: Belarinita, Yrjo
; TITLE OF INVENTION: Scarecrow Gene, Promoter and Uses Thereof
; FILE REFERENCE: 5914-075-999
; CURRENT APPLICATION NUMBER: US/09/186,276B
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 08/842,445
```

```

; PRIOR FILING DATE: 1997-04-24
; PRIOR APPLICATION NUMBER: 08/638,617
; PRIOR FILING DATE: 1996-04-26
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 59
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Oryza sp.
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(90)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-186-276B-59

Query Match
Best Local Similarity 100.0%; Score 18; DB 4; Length 90;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SIDA 4
   |||
Db 27 SIDA 30

RESULT 26
US-08-842-445-59
; Sequence 59, Application US/08842445A
; Patent No. 6441270
; GENERAL INFORMATION:
; APPLICANT: Benfey et al.
; TITLE OF INVENTION: Scarecrow Gene, Promoter and Uses
; FILE REFERENCE: 5914-056-999
; CURRENT APPLICATION NUMBER: US/08/842,445A
; CURRENT FILING DATE: 1997-04-24
; EARLIER APPLICATION NUMBER: 08/638,617
; EARLIER FILING DATE: 1996-04-26
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 59
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Plant
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(90)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-08-842-445-59

Query Match
Best Local Similarity 100.0%; Score 18; DB 4; Length 90;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SIDA 4
   |||
Db 27 SIDA 30

RESULT 27
US-09-186-188B-59
; Sequence 59, Application US/09186188B
; Patent No. 6455672
; GENERAL INFORMATION:
; APPLICANT: Benfey et al.
; TITLE OF INVENTION: Scarecrow Gene, Promoter and Uses
; FILE REFERENCE: 5914-074-999
; CURRENT APPLICATION NUMBER: US/09/186,188B
; CURRENT FILING DATE: 1998-11-05
; PRIOR FILING DATE: 1997-04-24
; PRIOR APPLICATION NUMBER: 08/842,445
; PRIOR FILING DATE: 1996-04-26
```

NUMBER OF SEQ ID NOS: 79  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 59  
LENGTH: 90  
TYPE: PRT  
ORGANISM: Plant  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (1)...(90)  
OTHER INFORMATION: Xaa - Any Amino Acid  
US-09-186-188B-59

Query Match  
Best Local Similarity 100.0%; Score 18; DB 4; Length 90;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
1111  
DB 27 SLDA 30

RESULT 28  
US-08-936-165A-515  
Sequence 515, Application US/08936165A  
Patent No. 6348582  
GENERAL INFORMATION:  
APPLICANT: Black, Michael  
APPLICANT: Burnham, Martin  
APPLICANT: Hodgson, John  
APPLICANT: Knowles, David  
APPLICANT: Lonetto, Michael  
APPLICANT: Nicholas, Richard  
APPLICANT: Pratt, Julie  
APPLICANT: Reichard, Richard  
APPLICANT: Rosenberg, Martin  
APPLICANT: Ward, Judith  
TITLE OF INVENTION: NO. 6348582el Prokaryotic Polynucleotides,  
TITLE OF INVENTION: Polypeptides and Their Uses  
NUMBER OF SEQUENCES: 534  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/936,165A  
FILING DATE: 24-SEP-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/027,032  
FILING DATE: 24-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimmil, Edward R  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P50549  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 515:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 105 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Protein

US-08-936-165A-515

Query Match  
Best Local Similarity 100.0%; Score 18; DB 4; Length 105;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
1111  
DB 4 SLDA 7

RESULT 29  
US-09-099-041A-6  
Sequence 6, Application US/09099041A  
Patent No. 6340576  
GENERAL INFORMATION:  
APPLICANT: Berlin, John  
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED  
TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF  
FILE REFERENCE: 07334-076001  
CURRENT APPLICATION NUMBER: US/09/099,041A  
CURRENT FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 09/019,942  
PRIOR FILING DATE: 1998-02-06  
NUMBER OF SEQ ID NOS: 37  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 109  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-099-041A-6

Query Match  
Best Local Similarity 100.0%; Score 18; DB 4; Length 109;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
1111  
DB 28 SLDA 31

RESULT 30  
US-09-199-637A-231  
Sequence 231, Application US/09199637A  
Patent No. 6355411  
GENERAL INFORMATION:  
APPLICANT: Ausubel, Frederick  
APPLICANT: Goodman, Howard M.  
APPLICANT: Rahme, Laurence G.  
APPLICANT: Mahajan-Miklos, Shalina  
APPLICANT: Tan, Man-Wan  
APPLICANT: Cao, Hui  
APPLICANT: Drenkard, Eliana  
APPLICANT: Tsongalis, John  
TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID  
TITLE OF INVENTION: SEQUENCES AND USES THEREOF  
FILE REFERENCE: 00786/361002  
CURRENT APPLICATION NUMBER: US/09/199,637A  
CURRENT FILING DATE: 1998-11-25  
PRIOR APPLICATION NUMBER: 60/066,517  
PRIOR FILING DATE: 1997-11-25  
NUMBER OF SEQ ID NOS: 437  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 231  
LENGTH: 109  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-199-637A-231

Query Match  
Best Local Similarity 100.0%; Score 18; DB 4; Length 109;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
||||  
Db 30 SLDA 33

RESULT 31  
US-09-245-281-6  
; Sequence 6, Application US/09245281  
; Patent No. 6369196  
; GENERAL INFORMATION:  
; APPLICANT: Bertin, John  
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY  
; FILE REFERENCE: 07334/118001  
; CURRENT APPLICATION NUMBER: US/09/245,281  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: US 09/207,359  
; EARLIER FILING DATE: 1998-12-08  
; EARLIER APPLICATION NUMBER: US 09/099,041  
; EARLIER FILING DATE: 1998-06-17  
; EARLIER APPLICATION NUMBER: US 09/019,942  
; EARLIER FILING DATE: 1998-02-06  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 109  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-245-281-6

Query Match 100.0%; Score 18; DB 4; Length 109;  
Best Local Similarity 100.0%; Pred. No. 5.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
||||  
Db 28 SLDA 31

RESULT 32  
US-09-207-359B-6  
; Sequence 6, Application US/09207359B  
; Patent No. 6469140  
; GENERAL INFORMATION:  
; APPLICANT: Bertin, John  
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED  
; FILE REFERENCE: 07334-112001  
; CURRENT APPLICATION NUMBER: US/09/207,359B  
; CURRENT FILING DATE: 1998-12-08  
; PRIOR APPLICATION NUMBER: US 09/099,041  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: US 09/019,942  
; PRIOR FILING DATE: 1998-02-06  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 110  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-207-359B-6

Query Match 100.0%; Score 18; DB 4; Length 110;  
Best Local Similarity 100.0%; Pred. No. 5.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
||||  
Db 29 SLDA 32

RESULT 33  
US-09-437-054A-2

; Sequence 2, Application US/09437054A  
; Patent No. 6316698  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Stephen M.  
; APPLICANT: Kinney, Anthony J.  
; TITLE OF INVENTION: Plant Alpha-Glucosidase II Homologs  
; FILE REFERENCE: B81273 US NA  
; CURRENT APPLICATION NUMBER: US/09/437,054A  
; CURRENT FILING DATE: 2001-05-14  
; PRIOR APPLICATION NUMBER: 60/107,909  
; PRIOR FILING DATE: 1998-No. 6316698member-10  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 2  
; LENGTH: 114  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (85)  
US-09-437-054A-2

Query Match 100.0%; Score 18; DB 4; Length 114;  
Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
||||  
Db 53 SLDA 56

RESULT 34  
US-08-470-670A-15  
; Sequence 15, Application US/08470670A  
; Patent No. 5834309  
; Patent No. 5834309 5710045  
; GENERAL INFORMATION:  
; APPLICANT: Thompson, Craig B. B.  
; TITLE OF INVENTION: VERTEBRATE APOPTOSIS GENE:  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P. O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: United States of America  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/470,670A  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/081,448  
; FILING DATE: 22-JUN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Highlander, Steven L.  
; REGISTRATION NUMBER: 37,642  
; REFERENCE/DOCKET NUMBER: ARCD-090--1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (512) 418-3000  
; TELEFAX: (512) 474-7577  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 121 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:

TOPOLOGY: linear  
US-08-470-670A-15

Query Match  
Best Local Similarity 100.0%; Score 18; DB 2; Length 121;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
1111  
Db 70 SLDA 73

RESULT 35  
US-08-461-511A-15  
Sequence 15, Application US/08461511A  
Patent No. 6303331

GENERAL INFORMATION:  
APPLICANT: Thompson, Craig B.B.  
Boise, Lawrence H.  
TITLE OF INVENTION: VERTEBRATE APOPTOSIS GENE: COMPOSITIONS  
AND METHODS  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: United States of America  
ZIP: 77210

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/461,511A  
FILING DATE: 05-Jun-1995  
CLASSIFICATION: UNKNOWN  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: ARCD:179  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (512) 474-7577

INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 121 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 15:  
US-08-461-511A-15

Query Match  
Best Local Similarity 100.0%; Score 18; DB 4; Length 121;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
1111  
Db 70 SLDA 73

RESULT 36  
US-09-134-001C-2929  
Sequence 2929, Application US/09134001C  
Patent No. 6380370

GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C

CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 2929  
LENGTH: 124  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-2929

Query Match  
Best Local Similarity 100.0%; Score 18; DB 4; Length 124;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
1111  
Db 115 SLDA 118

RESULT 37  
US-08-213-452-2  
Sequence 2, Application US/08213452  
Patent No. 5695956

GENERAL INFORMATION:  
APPLICANT: McClane, Bruce A.  
APPLICANT: Hanna, Philip C.  
TITLE OF INVENTION: Clostridium perfringens Type A  
TITLE OF INVENTION: Enterotoxin Toxoid and Methods of Preparation and Use As A  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lewis F. Gould, Jr.  
STREET: 1700 Market St., Suite 3232  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: US  
ZIP: 19103

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/213,452  
FILING DATE:  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Gould, Jr., Lewis F.  
REGISTRATION NUMBER: 25,057  
REFERENCE/DOCKET NUMBER: 107291-3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 575-6000  
TELEFAX: (215) 575-6015

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 149 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-213-452-2

Query Match  
Best Local Similarity 100.0%; Score 18; DB 1; Length 149;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
1111  
Db 120 SLDA 123

```
RESULT 38
US-08-606-143-18
; Sequence 18, Application US/08606143
; Patent No. 5856155
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; TITLE OF INVENTION: COMPOUNDS AND RELATED METHODS FOR
; TITLE OF INVENTION: MODULATING POTASSIUM ION CHANNELS AND ASSAYS FOR SUCH
; TITLE OF INVENTION: COMPOUNDS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Volt & Mayer, Ltd.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/606,143
; FILING DATE: 23-FEB-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kilyk Jr., John
; REGISTRATION NUMBER: 30763
; REFERENCE/DOCKET NUMBER: 71756
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5600
; TELEFAX: (312) 616-5700
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 152 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-606-143-18

Query Match
Best Local Similarity 100.0%; Score 18; DB 2; Length 152;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4
Db 53 SLDA 56

RESULT 39
US-08-954-395A-16
; Sequence 16, Application US/08954395A
; Patent No. 6204022
; GENERAL INFORMATION:
; APPLICANT: Johnson, Howard M.
; APPLICANT: Subramaniam, Prem S.
; APPLICANT: Pontzer, Carol H.
; APPLICANT: Villarete, Lorelie H.
; APPLICANT: Campos, Jackeline
; APPLICANT: Chung, Albert D.
; APPLICANT: Li, Wayne W.
; APPLICANT: Liu, Philip T.
; TITLE OF INVENTION: LOW-TOXICITY HUMAN INTERFERON-ALPHA
; TITLE OF INVENTION: ANALOG
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates, LLP
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
```

```

; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM compatible
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/954,395A
; FILING DATE: Filed Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/631,328
; FILING DATE: 12-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dehlinger, Peter J
; REGISTRATION NUMBER: 27008
; REFERENCE/DOCKET NUMBER: 5600-0001.35
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; TELEX:
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 166 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; CLONE: HuIFN-alpha analog IFN-alpha N6
US-08-954-395A-16

Query Match
Best Local Similarity 100.0%; Score 18; DB 4; Length 166;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4
Db 8 SLDA 11

RESULT 40
US-09-069-023-6
; Sequence 6, Application US/09069023A
; Patent No. 6348573
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; APPLICANT: Koseki, Takeyoshi
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
; TITLE OF INVENTION: SIGNALING PATHWAY INHIBITORS AND ACTIVATORS
; FILE REFERENCE: US-03333
; CURRENT APPLICATION NUMBER: US/09/069,023A
; CURRENT FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 167
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-069-023-6

Query Match
Best Local Similarity 100.0%; Score 18; DB 4; Length 167;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4
Db 86 SLDA 89

RESULT 41
```



US-08-081-448-8  
; Sequence 8, Application US/08081448  
; Patent No. 5646008  
; GENERAL INFORMATION:  
; APPLICANT: Thompson, Craig B.  
; APPLICANT: Boise, Lawrence H.  
; TITLE OF INVENTION: Vertebrate Apoptosis Gene:  
; TITLE OF INVENTION: Compositions and Methods  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: 321 No. 5646008th Clark Street, Suite 800  
; CITY: Chicago  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60610  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/081,448  
; FILING DATE: 19930622  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: NO. 5646008thrup, Thomas E.  
; REGISTRATION NUMBER: 33,268  
; REFERENCE/DOCKET NUMBER: ARCD090  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-744-0090  
; TELEFAX: 312-755-4489  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 170 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-081-448-8  
Query Match 100.0%; Score 18; DB 1; Length 170;  
Best Local Similarity 100.0%; Pred. No. 8.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 SLDA 4  
Db 74 SLDA 77  
RESULT 42  
US-08-470-670A-9  
; Sequence 9, Application US/08470670A  
; Patent No. 5834309  
; Patent No. 5834309 5710045  
; GENERAL INFORMATION:  
; APPLICANT: Thompson, Craig B. B.  
; APPLICANT: Boise, Lawrence H.  
; TITLE OF INVENTION: VERTEBRATE APOPTOSIS GENE:  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: United States of America  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/470,670A  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/081,448  
; FILING DATE: 22-JUN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Highlander, Steven L.  
; REGISTRATION NUMBER: 37,642  
; REFERENCE/DOCKET NUMBER: ARCD:090--1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (512) 418-3000  
; TELEFAX: (512) 474-7577  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 170 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-470-670A-9  
Query Match 100.0%; Score 18; DB 2; Length 170;  
Best Local Similarity 100.0%; Pred. No. 8.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 SLDA 4  
Db 74 SLDA 77  
RESULT 43  
US-08-461-511A-9  
; Sequence 9, Application US/08461511A  
; Patent No. 6303331  
; GENERAL INFORMATION:  
; APPLICANT: Thompson, Craig B. B.  
; APPLICANT: Boise, Lawrence H.  
; TITLE OF INVENTION: VERTEBRATE APOPTOSIS GENE: COMPOSITIONS  
; AND METHODS  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: United States of America  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/461,511A  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: UNKNOWN  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Highlander, Steven L.  
; REGISTRATION NUMBER: 37,642  
; REFERENCE/DOCKET NUMBER: ARCD:179  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (512) 418-3000  
; TELEFAX: (512) 474-7577  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 170 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 9:  
US-08-461-511A-9  
Query Match 100.0%; Score 18; DB 4; Length 170;

Best Local Similarity 100.0%; Pred. No. 8.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SIDA 4  
Db 74 SIDA 77

RESULT 44  
US-09-271-014A-8

; Sequence 8, Application US/09271014A

; Patent No. 6395510

; GENERAL INFORMATION:

; APPLICANT: THOMPSON, CRAIG B.

; APPLICANT: BOISE, LAWRENCE H.

; TITLE OF INVENTION: VERTEBRATE APOPTOSIS GENE: COMPOSITIONS AND METHODS

; FILE REFERENCE: ARCD:316

; CURRENT APPLICATION NUMBER: US/09/271,014A

; NUMBER OF FILING DATE: 1999-03-17

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 8

; LENGTH: 170

; TYPE: PRT

; ORGANISM: Human

US-09-271-014A-8

Query Match 100.0%; Score 18; DB 4; Length 170;

Best Local Similarity 100.0%; Pred. No. 8.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SIDA 4  
Db 74 SIDA 77

RESULT 45

PCT-US94-07089-9

; Sequence 9, Application PC/TUS9407089

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: Vertebrate Apoptosis Gene:

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; STREET: P.O. Box 4433

; CITY: Houston

; STATE: TX

; COUNTRY: United States of America

; ZIP: 77210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS, ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US94/07089

; FILING DATE: CONCURRENTLY FILED

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/081,448

; FILING DATE: 22 JUNE 1993

; ATTORNEY/AGENT INFORMATION:

; NAME: PARKER, David L.

; REGISTRATION NUMBER: 32,165

; TELEPHONE: 512-320-7200

; TELEFAX: 713-789-2679

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 170 amino acids

; TYPE: amino acid

; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US94-07089-9

Query Match

Best Local Similarity 100.0%; Score 18; DB 5; Length 170;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SIDA 4  
Db 74 SIDA 77

Search completed: February 6, 2003, 11:24:07  
Job time: 10.3333 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 11:22:09 ; Search time 5.83333 Seconds  
(without alignments)  
15.202 Million cell updates/sec

Title: PAT943-4  
Perfect score: 18  
Sequence: 1 sida 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 129505 seqs, 22169297 residues  
Total number of hits satisfying chosen parameters: 508

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubppaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/2/pubppaa/PC1\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubppaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubppaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubppaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubppaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubppaa/PC1US\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/2/pubppaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubppaa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubppaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/2/pubppaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/2/pubppaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/2/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	100.0	6	8	US-08-981-824-45
2	18	100.0	7	9	US-09-758-426-41
3	18	100.0	7	9	US-09-758-426-44
4	18	100.0	7	9	US-09-758-198-41
5	18	100.0	7	9	US-09-758-188-44
6	18	100.0	7	10	US-09-997-900-10
7	18	100.0	7	10	US-09-758-128-41
8	18	100.0	7	10	US-09-758-128-44
9	18	100.0	13	9	US-09-309-196-96
10	18	100.0	19	10	US-09-739-354-118
11	18	100.0	19	10	US-09-904-615-118
12	18	100.0	34	10	US-09-864-761-45603
13	18	100.0	44	10	US-09-864-761-35349
14	18	100.0	45	9	US-10-004-717-54
15	18	100.0	50	10	US-09-729-674-68
16	18	100.0	53	10	US-09-727-801-8
17	18	100.0	55	10	US-09-864-761-34730
18	18	100.0	58	9	US-10-001-876-181
19	18	100.0	61	10	US-09-764-887-267

20	18	100.0	62	10	US-09-864-761-46194	Sequence 46194, A
21	18	100.0	64	9	US-09-796-692-1225	Sequence 1225, Ap
22	18	100.0	70	10	US-09-864-761-34065	Sequence 34065, A
23	18	100.0	73	8	US-08-424-5508-33	Sequence 33, Appl
24	18	100.0	77	10	US-09-764-865-1073	Sequence 1073, Ap
25	18	100.0	77	10	US-09-925-300-1880	Sequence 1880, Ap
26	18	100.0	78	10	US-09-764-877-1348	Sequence 1348, Ap
27	18	100.0	79	10	US-09-864-761-40746	Sequence 40746, A
28	18	100.0	81	10	US-09-764-877-1555	Sequence 1555, Ap
29	18	100.0	83	10	US-09-925-300-1597	Sequence 1597, Ap
30	18	100.0	85	10	US-09-893-737-56	Sequence 56, Appl
31	18	100.0	86	9	US-10-108-605-281	Sequence 281, App
32	18	100.0	89	10	US-09-931-071-11	Sequence 11, Appl
33	18	100.0	90	10	US-09-841-8798-10	Sequence 10, Appl
34	18	100.0	92	9	US-10-002-974-21	Sequence 21, Appl
35	18	100.0	92	12	US-10-014-269-21	Sequence 21, Appl
36	18	100.0	96	10	US-09-867-550-338	Sequence 338, App
37	18	100.0	98	10	US-09-905-243-48	Sequence 48, Appl
38	18	100.0	98	10	US-09-905-243-49	Sequence 49, Appl
39	18	100.0	105	10	US-09-939-980-515	Sequence 515, App
40	18	100.0	109	9	US-10-118-984-6	Sequence 6, Appli
41	18	100.0	109	9	US-09-975-719-231	Sequence 231, App
42	18	100.0	109	10	US-09-728-721-6	Sequence 6, Appli
43	18	100.0	109	12	US-10-105-931-6	Sequence 6, Appli
44	18	100.0	118	10	US-09-216-393-58	Sequence 58, Appl
45	18	100.0	118	10	US-09-867-550-1788	Sequence 1788, Ap

#### ALIGNMENTS

RESULT 1  
US-08-981-824-45  
; Sequence 45, Application US/08981824  
; Patent No. US20020114816A1  
; GENERAL INFORMATION:  
; APPLICANT: ENDL, Josef  
; APPLICANT: STAHL, Peter  
; APPLICANT: ALBERT, Winfried  
; APPLICANT: SCHEDEL, Dolores  
; APPLICANT: BOITARD, Christian  
; APPLICANT: VAN ENDERT, Peter  
; APPLICANT: JUNG, Gunther-Gerhard  
; TITLE OF INVENTION: AUTOREACTIVE PEPTIDES FROM HUMAN GLUTAMIC ACID  
; TITLE OF INVENTION: DECARBOXYLASE (GAD)  
; FILE REFERENCE: 564-7029  
; CURRENT APPLICATION NUMBER: US/08/981,824  
; EARLIER FILING DATE: 1998-09-18  
; EARLIER APPLICATION NUMBER: PCT/EP96/03093  
; EARLIER FILING DATE: 1996-07-15  
; EARLIER APPLICATION NUMBER: DE/195 25 784.7  
; EARLIER FILING DATE: 1995-07-14  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 45  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-08-981-824-45

Query Match 100.0%; Score 18; DB 8; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.1e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SIDA 4  
Db 1 SIDA 4

RESULT 2  
US-09-758-426-41  
; Sequence 41, Application US/09758426  
; Patent No. US20020169116A1

```

; GENERAL INFORMATION:
; APPLICANT: KINGSTON, David J.
; APPLICANT: GERRATY, No. US20020169116Alman L.
; APPLICANT: WESTBROOK, Simon L.
; TITLE OF INVENTION: MODULATING THE ACTIVITY OF HORMONES OR THEIR RECEPTORS
; FILE REFERENCE: 016786/0214
; CURRENT APPLICATION NUMBER: US/09/758,426
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 09/194,218
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: AU PN9990
; PRIOR FILING DATE: 1996-05-22
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 41
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-758-426-41
```

```

Query Match          100.0%; Score 18; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 SLDA 4
        ||||
DB      2 SLDA 5
```

```

RESULT 3
US-09-758-426-44
; Sequence 44, Application US/09758426
; Patent No. US20020169116A1
; GENERAL INFORMATION:
; APPLICANT: KINGSTON, David J.
; APPLICANT: GERRATY, No. US20020169116Alman L.
; APPLICANT: WESTBROOK, Simon L.
; TITLE OF INVENTION: MODULATING THE ACTIVITY OF HORMONES OR THEIR RECEPTORS
; FILE REFERENCE: 016786/0214
; CURRENT APPLICATION NUMBER: US/09/758,426
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 09/194,218
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: AU PN9990
; PRIOR FILING DATE: 1996-05-22
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 44
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Rat
US-09-758-426-44
```

```

Query Match          100.0%; Score 18; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 SLDA 4
        ||||
DB      2 SLDA 5
```

```

RESULT 4
US-09-758-198-41
; Sequence 41, Application US/09758198
; Patent No. US20020187925A1
; GENERAL INFORMATION:
; APPLICANT: KINGSTON, David J.
; APPLICANT: GERRATY, No. US20020187925Alman L.
; APPLICANT: WESTBROOK, Simon L.
; TITLE OF INVENTION: MODULATING THE ACTIVITY OF HORMONES OR THEIR RECEPTORS
```

```

; TITLE OF INVENTION: - PEPTIDES, ANTIBODIES, VACCINES AND USES THEREOF
; FILE REFERENCE: 016786/0214
; CURRENT APPLICATION NUMBER: US/09/758,198
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/194,218
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: PCT/AU97/00312
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-05-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: AU PN9990
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-05-22
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 41
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-758-198-41
```

```

Query Match          100.0%; Score 18; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 SLDA 4
        ||||
DB      2 SLDA 5
```

```

RESULT 5
US-09-758-198-44
; Sequence 44, Application US/09758198
; Patent No. US20020187925A1
; GENERAL INFORMATION:
; APPLICANT: KINGSTON, David J.
; APPLICANT: GERRATY, No. US20020187925Alman L.
; APPLICANT: WESTBROOK, Simon L.
; TITLE OF INVENTION: MODULATING THE ACTIVITY OF HORMONES OR THEIR RECEPTORS
; FILE REFERENCE: 016786/0214
; CURRENT APPLICATION NUMBER: US/09/758,198
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/194,218
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: PCT/AU97/00312
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-05-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: AU PN9990
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-05-22
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 44
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Rat
US-09-758-198-44
```

```

Query Match          100.0%; Score 18; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 SLDA 4
        ||||
DB      2 SLDA 5
```

```

RESULT 6
US-09-997-900-10
; Sequence 10, Application US/09997900
; Patent No. US20020053098A1
; GENERAL INFORMATION:
; APPLICANT: Kakefuda, Genichi
; APPLICANT: Costello, Colleen
; APPLICANT: Sun, Ming
; APPLICANT: Hu, Weiming
; TITLE OF INVENTION: Genes and Vectors for Confering Herbicide Resistance
```

;; TITLE OF INVENTION: In Plants  
;; FILE REFERENCE: 043753/241148 (5849-20A)  
;; CURRENT APPLICATION NUMBER: US/09/997,900  
;; CURRENT FILING DATE: 2001-11-30  
;; PRIOR APPLICATION NUMBER: 60/106,239  
;; PRIOR FILING DATE: 1998-10-29  
;; PRIOR APPLICATION NUMBER: 09/426,568  
;; PRIOR FILING DATE: 1999-10-22  
;; NUMBER OF SEQ ID NOS: 11  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 10  
;; LENGTH: 7  
;; TYPE: PRP  
;; ORGANISM: Arabidopsis sp.  
;; NAME/KEY: PEPTIDE  
;; LOCATION: (1)..(7)  
;; OTHER INFORMATION: N-terminal sequence of AHAS small subunit peptide  
;; NAME/KEY: SITE  
;; LOCATION: (1)..(2)  
;; OTHER INFORMATION: Thrombin cleavage site  
US-09-997-900-10

Query Match 100.0%; Score 18; DB 10; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.1e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|||  
DB 2 SLDA 5

RESULT 7  
US-09-758-128-41  
;; Sequence 41, Application US/09758128  
;; Patent No. US20020107187A1  
;; GENERAL INFORMATION:  
;; APPLICANT: KINGSTON, David J.  
;; APPLICANT: GERRATT, No. US20020107187A1man L.  
;; APPLICANT: WESTBROOK, Simon L.  
;; TITLE OF INVENTION: MODULATING THE ACTIVITY OF HORMONES OR THEIR RECEPTORS  
;; FILE REFERENCE: 016786/0214  
;; CURRENT APPLICATION NUMBER: US/09/758,128  
;; CURRENT FILING DATE: 2001-01-12  
;; PRIOR APPLICATION NUMBER: 09/194,218  
;; PRIOR FILING DATE: 1999-02-05  
;; PRIOR APPLICATION NUMBER: AU P99990  
;; PRIOR FILING DATE: 1996-05-22  
;; NUMBER OF SEQ ID NOS: 58  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 41  
;; LENGTH: 7  
;; TYPE: PRP  
;; ORGANISM: Homo sapiens  
US-09-758-128-41

Query Match 100.0%; Score 18; DB 10; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.1e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|||  
DB 2 SLDA 5

RESULT 8  
US-09-758-128-44  
;; Sequence 44, Application US/09758128  
;; Patent No. US20020107187A1  
;; GENERAL INFORMATION:  
;; APPLICANT: KINGSTON, David J.

;; APPLICANT: GERRATT, No. US20020107187A1man L.  
;; APPLICANT: WESTBROOK, Simon L.  
;; TITLE OF INVENTION: MODULATING THE ACTIVITY OF HORMONES OR THEIR RECEPTORS  
;; FILE REFERENCE: 016786/0214  
;; CURRENT APPLICATION NUMBER: US/09/758,128  
;; CURRENT FILING DATE: 2001-01-12  
;; PRIOR APPLICATION NUMBER: 09/194,218  
;; PRIOR FILING DATE: 1999-02-05  
;; PRIOR APPLICATION NUMBER: AU P99990  
;; PRIOR FILING DATE: 1996-05-22  
;; NUMBER OF SEQ ID NOS: 58  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 44  
;; LENGTH: 7  
;; TYPE: PRP  
;; ORGANISM: Rat  
US-09-758-128-44

Query Match 100.0%; Score 18; DB 10; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.1e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|||  
DB 2 SLDA 5

RESULT 9  
US-09-309-196-96  
;; Sequence 96, Application US/09309196  
;; Publication NO. US20030008380A1  
;; GENERAL INFORMATION:  
;; APPLICANT: FOWLES, Dana M.  
;; APPLICANT: BROACH, Jim  
;; APPLICANT: MANFREDI, John  
;; APPLICANT: KLEIN, Christine  
;; APPLICANT: MURPHY, Andrew J.  
;; APPLICANT: PAUL, Jeremy  
;; APPLICANT: TRUEHEART, Joshua  
;; TITLE OF INVENTION: YEAST CELLS ENGINEERED TO PRODUCE  
;; NUMBER OF SEQUENCES: 119  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: BROWDY AND NEWMARK  
;; STREET: 419 Seventh Street, N.W., Suite 300  
;; CITY: Washington  
;; STATE: D.C.  
;; COUNTRY: USA  
;; ZIP: 20004  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/309,196  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/322,137  
;; FILING DATE:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/190,328  
;; FILING DATE: 31-JAN-1994  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/041,431  
;; FILING DATE: 31-MAR-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: COOPER, Iver P.  
;; REGISTRATION NUMBER: 28,005  
;; REFERENCE/DOCKET NUMBER: FOLWKS-2C  
;; TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 96:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-309-196-96

Query Match 100.0%; Score 18; DB 9; Length 13;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
||||  
DB 5 SLDA 8

RESULT 10  
US-09-739-254-118  
; Sequence 118, Application US/09739254  
; Patent No. US20010021700A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 49 Human Secreted Proteins  
; FILE REFERENCE: P2032P1  
; CURRENT APPLICATION NUMBER: US/09/739,254  
; EARLIER FILING DATE: 2000-12-19  
; EARLIER APPLICATION NUMBER: 09/511,554  
; EARLIER FILING DATE: 2000-02-23  
; EARLIER APPLICATION NUMBER: PCT/US99/19330  
; EARLIER FILING DATE: 1999-08-24  
; EARLIER APPLICATION NUMBER: 60/097,917  
; EARLIER FILING DATE: 1998-08-25  
; EARLIER APPLICATION NUMBER: 60/098,634  
; EARLIER FILING DATE: 1998-08-31  
; NUMBER OF SEQ ID NOS: 170  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 118  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (19)  
; OTHER INFORMATION: Xaa equals stop translation  
US-09-739-254-118

Query Match 100.0%; Score 18; DB 10; Length 19;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
||||  
DB 1 SLDA 4

RESULT 11  
US-09-904-615-118  
; Sequence 118, Application US/09904615  
; Patent No. US20020026040A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 49 Human Secreted Proteins  
; FILE REFERENCE: P2032P1  
; CURRENT APPLICATION NUMBER: US/09/904,615  
; EARLIER FILING DATE: 2001-07-16  
; EARLIER APPLICATION NUMBER: 09/511,554  
; EARLIER FILING DATE: 2000-02-23  
; EARLIER APPLICATION NUMBER: 60/097,917  
; EARLIER FILING DATE: 1998-08-25

PRIOR APPLICATION NUMBER: 60/098,634  
PRIOR FILING DATE: 1998-08-31  
NUMBER OF SEQ ID NOS: 170  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 118  
LENGTH: 19  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (19)  
OTHER INFORMATION: Xaa equals stop translation  
US-09-904-615-118

Query Match 100.0%; Score 18; DB 10; Length 19;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
||||  
DB 1 SLDA 4

RESULT 12  
US-09-864-761-45603  
; Sequence 45603, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecm1ca-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; EARLIER FILING DATE: 2001-05-23  
; EARLIER APPLICATION NUMBER: US 60/180,312  
; EARLIER FILING DATE: 2000-02-04  
; EARLIER APPLICATION NUMBER: US 60/207,456  
; EARLIER FILING DATE: 2000-05-26  
; EARLIER APPLICATION NUMBER: US 09/632,366  
; EARLIER FILING DATE: 2000-08-03  
; EARLIER APPLICATION NUMBER: GB 24263,6  
; EARLIER FILING DATE: 2000-10-04  
; EARLIER APPLICATION NUMBER: US 60/236,359  
; EARLIER FILING DATE: 2000-09-27  
; EARLIER APPLICATION NUMBER: PCT/US01/00666  
; EARLIER FILING DATE: 2001-01-30  
; EARLIER APPLICATION NUMBER: PCT/US01/00667  
; EARLIER FILING DATE: 2001-01-30  
; EARLIER APPLICATION NUMBER: PCT/US01/00664  
; EARLIER FILING DATE: 2001-01-30  
; EARLIER APPLICATION NUMBER: PCT/US01/00669  
; EARLIER FILING DATE: 2001-01-30  
; EARLIER APPLICATION NUMBER: PCT/US01/00665  
; EARLIER FILING DATE: 2001-01-30  
; EARLIER APPLICATION NUMBER: PCT/US01/00661  
; EARLIER FILING DATE: 2001-01-30  
; EARLIER APPLICATION NUMBER: PCT/US01/00670  
; EARLIER FILING DATE: 2001-01-30  
; EARLIER APPLICATION NUMBER: US 60/234,687  
; EARLIER FILING DATE: 2000-09-21  
; EARLIER APPLICATION NUMBER: US 09/608,408  
; EARLIER FILING DATE: 2000-06-30  
; EARLIER APPLICATION NUMBER: US 09/774,203  
; EARLIER FILING DATE: 2001-01-29



APPLICANT: Jacobs, Kenneth  
 APPLICANT: McCoy, John M.  
 APPLICANT: Lavallee, Edward R.  
 APPLICANT: Collins-Racie, Lisa A.  
 APPLICANT: Evans, Cheryl  
 APPLICANT: Merberg, David  
 APPLICANT: Treacy, Maurice  
 APPLICANT: Agostino, Michael J.  
 APPLICANT: Steindinger II, Robert J.  
 APPLICANT: Spaulding, Vikki  
 APPLICANT: Wong, Gordon G.  
 APPLICANT: Clark, Hilary  
 APPLICANT: Pechtel, Kim  
 APPLICANT: Genetics Institute, Inc.  
 TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM  
 FILE REFERENCE: 6055-64x  
 CURRENT APPLICATION NUMBER: US/09/729,674  
 CURRENT FILING DATE: 2000-12-04  
 PRIOR APPLICATION NUMBER: 09/539,330  
 PRIOR FILING DATE: 2000-03-30  
 NUMBER OF SEQ ID NOS: 283  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 68  
 LENGTH: 50  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: UNSURE  
 LOCATION: (29)  
 NAME/KEY: UNSURE  
 LOCATION: (39)  
 NAME/KEY: UNSURE  
 LOCATION: (45)  
 US-09-729-674-68  
 Query Match 100.0%; Score 18; DB 10; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 91;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SLDA 4  
 Db 8 SLDA 11  
 RESULT 16  
 US-09-727-801-8  
 Sequence 8, Application US/09727801  
 Patent No. US20010034059A1  
 GENERAL INFORMATION:  
 APPLICANT: Allen, Steve  
 APPLICANT: Helentjaris, Tim  
 TITLE OF INVENTION: Homologs of SCF Ubiquitin-Ligase Complex Component GRB1  
 FILE REFERENCE: BBI418 US NA  
 CURRENT APPLICATION NUMBER: US/09/727,801  
 CURRENT FILING DATE: 2000-12-01  
 PRIOR APPLICATION NUMBER: 60/170377  
 PRIOR FILING DATE: 1999-12-13  
 NUMBER OF SEQ ID NOS: 21  
 SOFTWARE: Microsoft Office 97  
 SEQ ID NO 8  
 LENGTH: 53  
 TYPE: PRT  
 ORGANISM: Oryza sativa  
 FEATURE:  
 NAME/KEY: UNSURE  
 LOCATION: (46)  
 US-09-727-801-8  
 Query Match 100.0%; Score 18; DB 10; Length 53;  
 Best Local Similarity 100.0%; Pred. No. 97;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SLDA 4

Db 8 SLDA 11  
 RESULT 17  
 US-09-864-761-34730  
 Sequence 34730, Application US/09864761  
 Patent No. US20020048763A1  
 GENERAL INFORMATION:  
 APPLICANT: Penn, Sharon G.  
 APPLICANT: Rank, David R.  
 APPLICANT: Hanzel, David K.  
 APPLICANT: Chen, Wensheng  
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 FILE REFERENCE: Aecm104-X-1  
 CURRENT APPLICATION NUMBER: US/09/864,761  
 CURRENT FILING DATE: 2001-05-23  
 PRIOR APPLICATION NUMBER: US 60/180,312  
 PRIOR FILING DATE: 2000-02-04  
 PRIOR APPLICATION NUMBER: US 60/207,456  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: US 09/632,366  
 PRIOR FILING DATE: 2000-08-03  
 PRIOR APPLICATION NUMBER: GB 24263.6  
 PRIOR FILING DATE: 2000-10-04  
 PRIOR APPLICATION NUMBER: US 60/236,359  
 PRIOR FILING DATE: 2000-09-27  
 PRIOR APPLICATION NUMBER: PCT/US01/00666  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00667  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00664  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00669  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00665  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00668  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00663  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00662  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00661  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00670  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: US 60/234,687  
 PRIOR FILING DATE: 2000-09-21  
 PRIOR APPLICATION NUMBER: US 09/608,408  
 PRIOR FILING DATE: 2000-06-30  
 PRIOR APPLICATION NUMBER: US 09/774,203  
 PRIOR FILING DATE: 2001-01-29  
 NUMBER OF SEQ ID NOS: 49117  
 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
 SEQ ID NO 34730  
 LENGTH: 55  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 OTHER INFORMATION: MAP TO AC009743.1  
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1  
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7  
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.9  
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7  
 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.4  
 OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.5  
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.1  
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2  
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6  
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.9  
 OTHER INFORMATION: SWISSPROT HIT: P09095, EVALU 3.40e+00



OTHER INFORMATION: EST\_HUMAN HIT: R09371.1, EVALUE 1.30e-02  
US-09-864-761-34730

Query Match 100.0%; Score 18; DB 10; Length 55;  
Best Local Similarity 100.0%; Pred. NO. 1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
DB 28 SLDA 31

RESULT 18  
US-10-001-876-181

Sequence 181, Application US/10001876  
Patent No. US20020177140A1  
GENERAL INFORMATION:  
APPLICANT: Salceda, Susana  
APPLICANT: Macina, Roberto  
APPLICANT: Recipon, Herve  
APPLICANT: Cafierkey, Robert  
APPLICANT: Ali, Shujath  
APPLICANT: Sun, Yongming  
APPLICANT: Liu, Chenghua  
TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and  
FILE REFERENCE: DEX-0285  
CURRENT APPLICATION NUMBER: US/10/001.876  
CURRENT FILING DATE: 2001-11-20  
PRIOR APPLICATION NUMBER: 60/252,186  
PRIOR FILING DATE: 2000-11-21  
NUMBER OF SEQ ID NOS: 211  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 181  
LENGTH: 58  
TYPE: PRT  
ORGANISM: Homo sapien  
US-10-001-876-181

Query Match 100.0%; Score 18; DB 9; Length 58;  
Best Local Similarity 100.0%; Pred. NO. 1.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
DB 25 SLDA 28

RESULT 19  
US-09-764-887-267

Sequence 267, Application US/09764887  
Patent No. US20020042096A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PALL3  
CURRENT APPLICATION NUMBER: US/09/764,887  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 658  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 267  
LENGTH: 61  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-764-887-267

Query Match 100.0%; Score 18; DB 10; Length 61;  
Best Local Similarity 100.0%; Pred. NO. 1.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
DB 17 SLDA 20

RESULT 20  
US-09-864-761-46194

Sequence 46194, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aecm1ca-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 46194  
LENGTH: 62  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC024049.2  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.6  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5  
OTHER INFORMATION: EST\_HUMAN HIT: AA868675.1, EVALUE 7.00e-31  
OTHER INFORMATION: SWISSPROT HIT: P08960, EVALUE 2.60e+00  
US-09-864-761-46194

Query Match 100.0%; Score 18; DB 10; Length 62;  
Best Local Similarity 100.0%; Pred. NO. 1.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4

Db 28 SLDA 31

## RESULT 21

US-09-796-692-1225  
; Sequence 1225, Application US/09796692  
; Publication No. US20020198362A1

## GENERAL INFORMATION:

APPLICANT: Galger, Alexander  
APPLICANT: Algate, Paul A.  
APPLICANT: Mannion, Jane

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
FILE REFERENCE: 2077.001200

CURRENT APPLICATION NUMBER: US/09/796,692

PRIOR FILING DATE: 2001-03-01

PRIOR APPLICATION NUMBER: 60/186,126

PRIOR FILING DATE: 2000-03-01

PRIOR APPLICATION NUMBER: 60/190,479

PRIOR FILING DATE: 2000-03-17

PRIOR APPLICATION NUMBER: 60/200,545

PRIOR FILING DATE: 2000-04-27

PRIOR APPLICATION NUMBER: 60/200,303

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: 60/200,779

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: 60/200,999

PRIOR FILING DATE: 2000-05-01

PRIOR APPLICATION NUMBER: 60/202,084

PRIOR FILING DATE: 2000-05-04

PRIOR APPLICATION NUMBER: 60/206,201

PRIOR FILING DATE: 2000-05-22

PRIOR APPLICATION NUMBER: 60/218,950

PRIOR FILING DATE: 2000-07-14

PRIOR APPLICATION NUMBER: 60/222,903

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: 60/223,416

PRIOR FILING DATE: 2000-08-04

PRIOR APPLICATION NUMBER: 60/223,378

PRIOR FILING DATE: 2000-08-07

NUMBER OF SEQ ID NOS: 9597

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 1225

LENGTH: 64

TYPE: prt

ORGANISM: Homo sapiens

US-09-796-692-1225

Qy 1 SLDA 4

Db 42 SLDA 45

## RESULT 22

US-09-864-761-34065  
; Sequence 34065, Application US/09864761  
; Patent No. US20020048763A1

## GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: Aecmica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: US 60/234,687

PRIOR FILING DATE: 2000-09-21

PRIOR APPLICATION NUMBER: US 09/608,408

PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: US 09/774,203

PRIOR FILING DATE: 2001-01-29

NUMBER OF SEQ ID NOS: 49117

SOFTWARE: Anomax Sequence Listing Engine vers. 1.1

SEQ ID NO 34065

LENGTH: 70

TYPE: prt

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO 283845.14

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.77

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1

OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8

OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.73

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4

OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.78

OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4

OTHER INFORMATION: EST\_HUMAN HIT: BE746542.1, EVALUE 5.00e-35

OTHER INFORMATION: SWISSPROT HIT: Q15750, EVALUE 4.00e-36

US-09-864-761-34065

## RESULT 23

US-08-424-550B-33  
; Sequence 33, Application US/08424550B  
; Patent No. US20020119447A1

## GENERAL INFORMATION:

APPLICANT: JOHN N. SIMONS  
APPLICANT: TAMI J. PILOTT-MATTIAS

Qy 1 SLDA 4

Db 11 SLDA 14

Query Match 100.0%; Score 18; DB 10; Length 70;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

APPLICANT: GEORGE J. DAMSON
APPLICANT: GEORGE G. SCHLAUDER
APPLICANT: SURESH M. DESAI
APPLICANT: THOMAS P. LEARY
APPLICANT: ANTHONY SCOTT MUEHRHOFF
APPLICANT: JAMES C. ERKER
APPLICANT: SHERI L. BUIK
APPLICANT: ISA K. MUSHAWAR
TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
NUMBER OF SEQUENCES: 716
CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
STREET: 100 ABBOTT PARK ROAD
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/424,550B
FILING DATE:
CLASSIFICATION: 435435
ATTORNEY/AGENT INFORMATION:
NAME: FOREMSKI, PRISCILLA E.
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 5527.PC.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 73 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-424-550B-33

Query Match 100.0%; Score 18; DB 8; Length 73;
Best Local Similarity 100.0%; Pred. NO. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY 1 SLDA 4
Db 58 SLDA 61

RESULT 24
US-09-764-869-1073
Sequence 1073, Application US/09764869
Patent No. US20020061521A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC007
CURRENT APPLICATION NUMBER: US/09/764,869
PRIORITY FILING DATE: 2001-01-17
Prior application data removed - refer to PAM or file wrapper
NUMBER OF SEQ ID NOS: 2442
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1073
LENGTH: 77
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (26)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

```

```

: NAME/KEY: SITE
: LOCATION: (36)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-869-1073

Query Match
Best Local Similarity 100.0%; Score 18; DB 10; Length 77;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4
|||
Db 48 SLDA 51

RESULT 25
US-09-925-300-1880
: Sequence 1880, Application US/09925300
: Patent No. US20020151681A1
: GENERAL INFORMATION:
: APPLICANT: Craig Rosen,
: APPLICANT: Steve Ruden
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
: FILE REFERENCE: PA101
: CURRENT APPLICATION NUMBER: US/09/925,300
: CURRENT FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: PCT/US00/05968
: PRIOR FILING DATE: 2000-03-08
: PRIOR APPLICATION NUMBER: 60/124,270
: PRIOR FILING DATE: 1999-03-12
: NUMBER OF SEQ ID NOS: 1880
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1880
: LENGTH: 77
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-925-300-1880

Query Match
Best Local Similarity 100.0%; Score 18; DB 10; Length 77;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4
|||
Db 67 SLDA 70

RESULT 26
US-09-764-877-1348
: Sequence 1348, Application US/09764877
: Patent No. US20020147140A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PC005
: CURRENT APPLICATION NUMBER: US/09/764,877
: CURRENT FILING DATE: 2001-01-17
: Prior application data removed - refer to PALM or file wrapper
: NUMBER OF SEQ ID NOS: 4031
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1348
: LENGTH: 78
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (28)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-877-1348

Query Match
Best Local Similarity 100.0%; Score 18; DB 10; Length 78;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 SLDA 4  
 ||||  
 Db 59 SLDA 62

## RESULT 27

US-09-864-761-40746  
 ; Sequence 40746, Application US/09864761  
 ; Patent No. US20020048763A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Penn, Sharron G.  
 ; APPLICANT: Rank, David R.  
 ; APPLICANT: Hanzel, David K.  
 ; APPLICANT: Chen, Wensheng  
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 ; FILE REFERENCE: Aecomica-X-1  
 ; CURRENT APPLICATION NUMBER: US/09/864,761  
 ; CURRENT FILING DATE: 2001-05-23  
 ; PRIOR APPLICATION NUMBER: US 60/180,312  
 ; PRIOR FILING DATE: 2000-02-04  
 ; PRIOR APPLICATION NUMBER: US 60/207,456  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: US 09/632,366  
 ; PRIOR FILING DATE: 2000-08-03  
 ; PRIOR APPLICATION NUMBER: GB 24263,6  
 ; PRIOR FILING DATE: 2000-10-04  
 ; PRIOR APPLICATION NUMBER: US 60/236,359  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: US 60/234,687  
 ; PRIOR FILING DATE: 2000-09-21  
 ; PRIOR APPLICATION NUMBER: US 09/608,408  
 ; PRIOR FILING DATE: 2000-06-30  
 ; PRIOR APPLICATION NUMBER: US 09/774,203  
 ; PRIOR FILING DATE: 2001-01-29  
 ; NUMBER OF SEQ ID NOS: 49117  
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
 ; SEQ ID NO 40746  
 ; LENGTH: 79  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: MAP TO AL138796.2  
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.53  
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.86  
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.64  
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.54  
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.55  
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.6  
 ; OTHER INFORMATION: EST\_HUMAN HIT: H29179.1, EVALUE 2.00e-21  
 ; OTHER INFORMATION: SWISSPROT HIT: P21179, EVALUE 3.30e+00  
 ; US-09-864-761-40746

Query Match 100.0%; Score 18; DB 10; Length 79;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
 ||||  
 Db 2 SLDA 5

## RESULT 28

US-09-764-877-1555  
 ; Sequence 1555, Application US/09764877  
 ; Patent No. US20020147140A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; FILE REFERENCE: PC005  
 ; CURRENT APPLICATION NUMBER: US/09/764,877  
 ; CURRENT FILING DATE: 2001-01-17  
 ; Prior application data removed - refer to PAM or file wrapper  
 ; NUMBER OF SEQ ID NOS: 4031  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 1555  
 ; LENGTH: 81  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: SITE  
 ; LOCATION: (54)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; NAME/KEY: SITE  
 ; LOCATION: (63)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; NAME/KEY: SITE  
 ; LOCATION: (65)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; NAME/KEY: SITE  
 ; LOCATION: (68)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; NAME/KEY: SITE  
 ; LOCATION: (72)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; US-09-764-877-1555  
 ; Query Match 100.0%; Score 18; DB 10; Length 81;  
 ; Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 ; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
 ||||  
 Db 25 SLDA 28

## RESULT 29

US-09-925-300-1597  
 ; Sequence 1597, Application US/09925300  
 ; Patent No. US20020151681A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Craig Rosen,  
 ; APPLICANT: Steve Ruben  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
 ; FILE REFERENCE: PA101  
 ; CURRENT APPLICATION NUMBER: US/09/925,300  
 ; CURRENT FILING DATE: 2001-08-10  
 ; PRIOR APPLICATION NUMBER: PCT/US00/05988  
 ; PRIOR FILING DATE: 2000-03-08  
 ; PRIOR APPLICATION NUMBER: 60/124,270  
 ; PRIOR FILING DATE: 1999-03-12  
 ; NUMBER OF SEQ ID NOS: 1890  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 1597  
 ; LENGTH: 83  
 ; TYPE: PRT

ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (71)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-300-1597

Query Match 100.0%; Score 18; DB 10; Length 83;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
1111  
DB 15 SLDA 18

RESULT 30  
US-09-893-737-56  
Sequence 56, Application US/09893737  
Patent No. US20020110855A1  
GENERAL INFORMATION:  
APPLICANT: Sheppard, Paul O.  
APPLICANT: Presnell, Scott R.  
TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS  
FILE REFERENCE: 00-41  
CURRENT APPLICATION NUMBER: US/09/893,737  
CURRENT FILING DATE: 2001-06-28  
PRIOR APPLICATION NUMBER: US 60/215,446  
PRIOR FILING DATE: 2000-06-30  
NUMBER OF SEQ ID NOS: 329  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 56  
LENGTH: 85  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-893-737-56

Query Match 100.0%; Score 18; DB 10; Length 85;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
1111  
DB 60 SLDA 63

RESULT 31  
US-10-108-605-281  
Sequence 281, Application US/10108605  
Patent No. US20020160934A1  
GENERAL INFORMATION:  
APPLICANT: Broadus, Julie  
APPLICANT: Stam, Lynn  
APPLICANT: Bachmann, Jane  
APPLICANT: Kamdar, Kim  
TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE  
TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF  
FILE REFERENCE: 31133B  
CURRENT APPLICATION NUMBER: US/10/108,605  
CURRENT FILING DATE: 2002-03-27  
PRIOR APPLICATION NUMBER: US 09/761,142  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/176,418  
PRIOR FILING DATE: 2000-01-14  
NUMBER OF SEQ ID NOS: 361  
SOFTWARE: Patent Ver. 2.1  
SEQ ID NO 281  
LENGTH: 86  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
US-10-108-605-281

Query Match 100.0%; Score 18; DB 9; Length 86;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
1111  
DB 3 SLDA 6

RESULT 32  
US-09-931-071-11  
Sequence 11, Application US/09931071  
Patent No. US20020128219A1  
GENERAL INFORMATION:  
APPLICANT: Bertin, John  
APPLICANT: Alnemri, Emad S.  
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED  
FILE REFERENCE: 07334-335001  
CURRENT APPLICATION NUMBER: US/09/931,071  
CURRENT FILING DATE: 2002-03-18  
PRIOR APPLICATION NUMBER: 09/428,252  
PRIOR FILING DATE: 1999-10-27  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 11  
LENGTH: 89  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-931-071-11

Query Match 100.0%; Score 18; DB 10; Length 89;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
1111  
DB 25 SLDA 28

RESULT 33  
US-09-841-879B-10  
Sequence 10, Application US/09841879B  
Patent No. US20020142979A1  
GENERAL INFORMATION:  
APPLICANT: Bertin, John  
FILE REFERENCE: 07334-330001  
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES TH  
CURRENT APPLICATION NUMBER: US/09/841,879B  
CURRENT FILING DATE: 2001-04-24  
PRIOR APPLICATION NUMBER: US 09/728,721  
PRIOR FILING DATE: 2000-12-01  
PRIOR APPLICATION NUMBER: US 09/340,620  
PRIOR FILING DATE: 1999-06-28  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10  
LENGTH: 90  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-841-879B-10

Query Match 100.0%; Score 18; DB 10; Length 90;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
1111  
DB 25 SLDA 28

RESULT 34  
US-10-002-974-21  
Sequence 21, Application US/10002974

```
; Publication No. US20020197616A1
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; APPLICANT: Ogura, Yasunori
; APPLICANT: Cho, Judy
; APPLICANT: Nicolae, Dan L
; APPLICANT: Bonen, Denise
; TITLE OF INVENTION: NOD2 Nucleic Acids and Proteins
; FILE REFERENCE: UM-06646
; CURRENT APPLICATION NUMBER: US/10/002,974
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-002-974-21
```

```
Query Match
Best Local Similarity 100.0%; Score 18; DB 9; Length 92;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SLDA 4
    ||||
Db 26 SLDA 29
```

```
RESULT 35
US-10-014-269-21
; Sequence 21, Application US/10014269
; Patent No. US20020127673A1
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; APPLICANT: Ogura, Yasunori
; TITLE OF INVENTION: NOD2 Nucleic Acids and Proteins
; FILE REFERENCE: UM-06645
; CURRENT APPLICATION NUMBER: US/10/014,269
; CURRENT FILING DATE: 2001-10-26
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-014-269-21
```

```
Query Match
Best Local Similarity 100.0%; Score 18; DB 12; Length 92;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SLDA 4
    ||||
Db 26 SLDA 29
```

```
RESULT 36
US-09-867-550-338
; Sequence 338, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehrahan, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
```

```
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 338
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)
; OTHER INFORMATION: wherein Xaa may be any one of Phe or Ser or Tyr or Cys or Leu
; OTHER INFORMATION: His or Arg or Ile or Thr or Asn or Val or Ala or Asp or Gly
US-09-867-550-338
```

```
Query Match
Best Local Similarity 100.0%; Score 18; DB 10; Length 96;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SLDA 4
    ||||
Db 51 SLDA 54
```

```
RESULT 37
US-09-905-243-48
; Sequence 48, Application US/09905243
; Patent No. US20020062009A1
; GENERAL INFORMATION:
; APPLICANT: Taylor, Alexander H
; TITLE OF INVENTION: Monoclonal Antibodies with Reduced
; TITLE OF INVENTION: Immunogenicity
; FILE REFERENCE: P50770
; CURRENT APPLICATION NUMBER: US/09/905,243
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/300,970
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 48
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Macaca cynomolgus
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (31)...(35)
; OTHER INFORMATION: CDRI
; NAME/KEY: DOMAIN
; LOCATION: (50)...(66)
; OTHER INFORMATION: CDRII
US-09-905-243-48
```

```
Query Match
Best Local Similarity 100.0%; Score 18; DB 10; Length 98;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SLDA 4
    ||||
Db 71 SLDA 74
```

```
RESULT 38
US-09-905-243-49
; Sequence 49, Application US/09905243
; Patent No. US20020062009A1
; GENERAL INFORMATION:
; APPLICANT: Taylor, Alexander H
; TITLE OF INVENTION: Monoclonal Antibodies with Reduced
; TITLE OF INVENTION: Immunogenicity
; FILE REFERENCE: P50770
; CURRENT APPLICATION NUMBER: US/09/905,243
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/300,970
```

PRIOR FILING DATE: 1999-04-28  
NUMBER OF SEQ ID NOS: 97  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 49  
LENGTH: 98  
TYPE: PRT  
ORGANISM: Macaca cynomolgus  
FEATURE:  
NAME/KEY: DOMAIN  
LOCATION: (31)...(35)  
OTHER INFORMATION: CDRI  
NAME/KEY: DOMAIN  
LOCATION: (50)...(66)  
OTHER INFORMATION: CDRII  
US-09-905-243-49

Query Match  
Best Local Similarity 100.0%; Score 18; DB 10; Length 98;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
Db 71 SLDA 74

RESULT 39  
US-09-939-980-515  
Sequence 515, Application US/09939980  
Patent No. US20020082234A1  
GENERAL INFORMATION:  
APPLICANT: Black, Michael  
Burnham, Martin  
Hodgson, John  
Knowles, David  
Lometto, Michael  
Nicholas, Richard  
Pratt, Julie  
Reichard, Richard  
Rosenberg, Martin  
Ward, Judith

TITLE OF INVENTION: NO. US20020082234A1el Prokaryotic Polynucleotides,  
Polypeptides and Their Uses  
NUMBER OF SEQUENCES: 534  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/939,980  
FILING DATE: 27-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/936,165  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimmil, Edward R  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P50549  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 515:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 105 amino acids

TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 515:  
US-09-939-980-515

Query Match  
Best Local Similarity 100.0%; Score 18; DB 10; Length 105;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
Db 4 SLDA 7

RESULT 40  
US-10-118-984-6  
Sequence 6, Application US/10118984  
Publication No. US20020197693A1  
GENERAL INFORMATION:  
APPLICANT: Berlin, John  
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY  
FILE REFERENCE: 07334/118001  
CURRENT APPLICATION NUMBER: US/10/118,984  
CURRENT FILING DATE: 2002-04-09  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/245,281  
PRIOR FILING DATE: EARLIER FILING DATE: 1999-02-05  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 09/207,359  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-08  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 09/099,041  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 09/019,942  
NUMBER OF SEQ ID NOS: 44  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 109  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-118-984-6

Query Match  
Best Local Similarity 100.0%; Score 18; DB 9; Length 109;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
Db 28 SLDA 31

RESULT 41  
US-09-975-719-231  
Sequence 231, Application US/09975719  
Publication No. US20030022349A1  
GENERAL INFORMATION:  
APPLICANT: Ausubel, Frederick M.  
TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID  
FILE REFERENCE: 00786/361003  
CURRENT APPLICATION NUMBER: US/09/975,719  
CURRENT FILING DATE: 2001-10-10  
PRIOR APPLICATION NUMBER: US 09/199,637  
PRIOR FILING DATE: 1998-11-25  
PRIOR APPLICATION NUMBER: US 60/066,517  
PRIOR FILING DATE: 1997-11-25  
NUMBER OF SEQ ID NOS: 437  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 231  
LENGTH: 109  
TYPE: PRT

ORGANISM: Pseudomonas aeruginosa  
US-09-975-719-231

Query Match 100.0%; Score 18; DB 9; Length 109;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SIDA 4  
||||  
Db 30 SLDA 33

RESULT 42  
US-09-728-721-6

Sequence 6, Application US/09728721  
Patent No. US20020061845A1

GENERAL INFORMATION:

APPLICANT: Bertin, John

TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF

FILE REFERENCE: 0734-124001

CURRENT APPLICATION NUMBER: US/09/728,721

CURRENT FILING DATE: 2000-12-01

PRIOR APPLICATION NUMBER: 09/340,620

PRIOR FILING DATE: 1999-06-28

PRIOR APPLICATION NUMBER: US 09/207,359

PRIOR FILING DATE: 1998-12-08

PRIOR APPLICATION NUMBER: US 09/099,041

PRIOR FILING DATE: 1998-06-17

PRIOR APPLICATION NUMBER: US 09/019,942

PRIOR FILING DATE: 1998-02-06

NUMBER OF SEQ ID NOS: 71

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 6

LENGTH: 109

TYPE: PRT

ORGANISM: Homo sapiens

US-09-728-721-6

Query Match 100.0%; Score 18; DB 10; Length 109;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SIDA 4  
||||  
Db 28 SLDA 31

RESULT 43  
US-10-105-931-6

Sequence 6, Application US/10105931  
Patent No. US20020150987A1

GENERAL INFORMATION:

APPLICANT: Bertin, John

TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF

FILE REFERENCE: 0734-076001

CURRENT APPLICATION NUMBER: US/10/105,931

CURRENT FILING DATE: 2002-03-25

PRIOR APPLICATION NUMBER: 09/099,041

PRIOR FILING DATE: 1998-06-17

PRIOR APPLICATION NUMBER: 09/019,942

PRIOR FILING DATE: 1998-02-06

NUMBER OF SEQ ID NOS: 37

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 6

LENGTH: 109

TYPE: PRT

ORGANISM: Homo sapiens

US-10-105-931-6

Query Match 100.0%; Score 18; DB 12; Length 109;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SIDA 4  
||||  
Db 28 SLDA 31

RESULT 44  
US-09-216-393-58

Sequence 58, Application US/09216393  
Patent No. US20010014447A1

GENERAL INFORMATION:

APPLICANT: Milhaussen, Michael James

TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND

FILE REFERENCE: TX-1-C2

CURRENT APPLICATION NUMBER: US/09/216,393

CURRENT FILING DATE: 1998-12-18

EARLIER APPLICATION NUMBER: 08/994,825

EARLIER FILING DATE: 1997-12-19

NUMBER OF SEQ ID NOS: 364

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 58

LENGTH: 118

TYPE: PRT

ORGANISM: Toxoplasma gondii

US-09-216-393-58

Query Match 100.0%; Score 18; DB 10; Length 118;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SIDA 4  
||||  
Db 113 SLDA 116

RESULT 45  
US-09-867-550-1788

Sequence 1788, Application US/09867550  
Patent No. US20020082206A1

GENERAL INFORMATION:

APPLICANT: Leach, Martin D.

APPLICANT: Mehraban, Fuad

APPLICANT: Conley, Pamela

APPLICANT: Law, Debbie

APPLICANT: Topper, James

TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells a

FILE REFERENCE: 21402-013 (Cura-313)

CURRENT APPLICATION NUMBER: US/09/867,550

CURRENT FILING DATE: 2001-09-20

PRIOR APPLICATION NUMBER: USSN 60/208,427

PRIOR FILING DATE: 2000-05-30

NUMBER OF SEQ ID NOS: 2125

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1788

LENGTH: 118

TYPE: PRT

ORGANISM: Homo sapiens

US-09-867-550-1788

Query Match 100.0%; Score 18; DB 10; Length 118;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SIDA 4  
||||  
Db 110 SLDA 113

Search completed: February 6, 2003, 11:39:45  
Job time : 6.83333 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: February 6, 2003, 11:16:59 : Search time 133.167 Seconds  
(without alignments)  
19.366 Million cell updates/sec

Title: PAT943-4  
Perfect score: 18  
Sequence: 1 slide 4

Scoring table:  
BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 13114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

Pending Patents-AA\_Main:\*

1: /cgn2\_6/ptodata/1/paa/PCTUS\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep:\*  
7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep:\*  
8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep:\*  
9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep:\*  
10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep:\*  
11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep:\*  
12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep:\*  
13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep:\*  
14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep:\*  
15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep:\*  
16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep:\*  
17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep:\*  
18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep:\*  
19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep:\*  
20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep:\*  
21: /cgn2\_6/ptodata/1/paa/US097\_COMB.pep:\*  
22: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep:\*  
23: /cgn2\_6/ptodata/1/paa/US099\_COMB.pep:\*  
24: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep:\*  
25: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep:\*  
26: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep:\*  
27: /cgn2\_6/ptodata/1/paa/US60\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	100.0	6	13	US-08-981-824-45
2	18	100.0	7	1	PCT-US94-11416-48
3	18	100.0	7	8	US-08-462-860-48
4	18	100.0	7	8	US-08-462-860A-48
5	18	100.0	7	15	US-09-194-218-41
6	18	100.0	7	15	US-09-194-218-44

7	18	100.0	7	17	US-09-396-347B-53	Sequence 53, Appl
8	18	100.0	7	17	US-09-396-347D-53	Sequence 53, Appl
9	18	100.0	7	17	US-09-396-347E-53	Sequence 53, Appl
10	18	100.0	7	17	US-09-396-347F-53	Sequence 53, Appl
11	18	100.0	7	21	US-09-758-128-41	Sequence 41, Appl
12	18	100.0	7	21	US-09-758-128-44	Sequence 44, Appl
13	18	100.0	7	21	US-09-758-198-41	Sequence 41, Appl
14	18	100.0	7	21	US-09-758-198-44	Sequence 44, Appl
15	18	100.0	7	21	US-09-758-426-41	Sequence 41, Appl
16	18	100.0	7	21	US-09-758-426-44	Sequence 44, Appl
17	18	100.0	7	22	US-09-861-661-41	Sequence 41, Appl
18	18	100.0	7	22	US-09-861-661-44	Sequence 44, Appl
19	18	100.0	7	23	US-09-997-900-10	Sequence 10, Appl
20	18	100.0	8	9	US-08-526-708-12	Sequence 12, Appl
21	18	100.0	8	12	US-08-813-273-12	Sequence 12, Appl
22	18	100.0	8	16	US-09-228-866-12	Sequence 12, Appl
23	18	100.0	8	21	US-09-700-993-25	Sequence 25, Appl
24	18	100.0	8	23	US-09-922-227-12	Sequence 12, Appl
25	18	100.0	8	23	US-09-988-493-271	Sequence 271, Appl
26	18	100.0	8	24	US-10-014-340-647	Sequence 647, Appl
27	18	100.0	10	1	PCT-US97-21437-13	Sequence 13, Appl
28	18	100.0	10	5	US-09-109-836-13	Sequence 13, Appl
29	18	100.0	12	5	US-08-188-354-2	Sequence 2, Appl
30	18	100.0	12	22	US-09-831-253A-3	Sequence 3, Appl
31	18	100.0	13	8	US-08-461-383-96	Sequence 96, Appl
32	18	100.0	13	8	US-08-463-181-96	Sequence 96, Appl
33	18	100.0	13	10	US-08-689-172A-32	Sequence 32, Appl
34	18	100.0	13	10	US-08-689-172C-32	Sequence 32, Appl
35	18	100.0	13	16	US-09-286-166-96	Sequence 96, Appl
36	18	100.0	13	17	US-09-309-196-96	Sequence 96, Appl
37	18	100.0	13	21	US-09-747-774A-32	Sequence 32, Appl
38	18	100.0	13	23	US-09-953-354-32	Sequence 32, Appl
39	18	100.0	13	26	US-10-263-341-96	Sequence 96, Appl
40	18	100.0	14	22	US-09-831-253A-6	Sequence 96, Appl
41	18	100.0	15	8	US-08-400-306-1	Sequence 1, Appl
42	18	100.0	16	1	PCT-US99-28465-15	Sequence 15, Appl
43	18	100.0	16	18	US-09-461-061-15	Sequence 15, Appl
44	18	100.0	16	18	US-09-461-061A-15	Sequence 15, Appl
45	18	100.0	17	1	PCT-US02-09671-65	Sequence 65, Appl

## ALIGNMENTS

RESULT 1  
US-08-981-824-45  
Sequence 45, Application US/08981824  
GENERAL INFORMATION:  
APPLICANT: ENDL, Josef  
APPLICANT: STAHL, Peter  
APPLICANT: ALBERT, Winfried  
APPLICANT: SCHENDEL, Dolores  
APPLICANT: BOITARD, Christian  
APPLICANT: VAN ENDERT, Peter  
APPLICANT: JUNG, Gunther-Gerhard  
TITLE OF INVENTION: AUTOREACTIVE PEPTIDES FROM HUMAN GLUTAMIC ACID  
TITLE OF INVENTION: DECARBOXYLASE (GAD)  
FILE REFERENCE: 564-7029  
CURRENT APPLICATION NUMBER: US/08/981,824  
CURRENT FILING DATE: 1998-09-18  
EARLIER APPLICATION NUMBER: PCT/EP96/03093  
EARLIER FILING DATE: 1996-07-15  
EARLIER APPLICATION NUMBER: DE/195 25 784.7  
EARLIER FILING DATE: 1995-07-14  
NUMBER OF SEQ ID NOS: 47  
SOFTWARE: Patent Ver. 2.1  
SEQ ID NO 45  
LENGTH: 6  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-08-981-824-45  
Query Match 100.0%; Score 18; DB 13; Length 6;

Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
||||  
Db 1 SLDA 4

RESULT 2  
PCT-US94-11416-48

; Sequence 48, Application PC/TUS9411416

; GENERAL INFORMATION:

APPLICANT: The University of Pennsylvania

TITLE OF INVENTION: "Plasmodium falciparum Ribonucleotide Reductase, DNA

NUMBER OF SEQUENCES: 67

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seidel, Gonda, Lavorigna & Monaco, P.C.

STREET: Suite 1800

CITY: Philadelphia

STATE: Pennsylvania

COUNTRY: U.S.A.

ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 720 KB

COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS

SOFTWARE: Wordperfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/11416

FILING DATE: Not Yet Assigned

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/136,743

FILING DATE: 14 October 1993

ATTORNEY/AGENT INFORMATION:

NAME: Monaco, Daniel A.

REGISTRATION NUMBER: 30,480

REFERENCE/DOCKET NUMBER: 3957-10 PC

TELECOMMUNICATION INFORMATION:

TELEPHONE: (215) 568-8383

TELEFAX: (215) 568-5549

TELEX: None

INFORMATION FOR SEQ ID NO: 48:

SEQUENCE CHARACTERISTICS:

LENGTH: 7 amino acids

TYPE: amino acid

TOPOLOGY: linear

PCT-US94-11416-48

Query Match

Best Local Similarity 100.0%; Score 18; DB 1; Length 7;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
||||  
Db 2 SLDA 5

RESULT 3  
US-08-462-860-48

; Sequence 48, Application US/08462860

; GENERAL INFORMATION:

APPLICANT: Barry S. Cooperman, Harvey Rubin,  
APPLICANT: Jerome Salem, and Allison L. Fisher  
TITLE OF INVENTION: "Peptide Inhibitors of  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seidel, Gonda, Lavorigna & Monaco, P.C.  
STREET: Suite 1800, Two Penn Center Plaza  
CITY: Philadelphia

STATE: Pennsylvania

COUNTRY: U.S.A.

ZIP: 19102

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 720 KB

COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS

SOFTWARE: Wordperfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/462,860

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/136,743

FILING DATE: October 14, 1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Monaco, Daniel A.

REGISTRATION NUMBER: 30,480

REFERENCE/DOCKET NUMBER: 3957-10 D11

TELECOMMUNICATION INFORMATION:

TELEPHONE: (215) 568-8383

TELEFAX: (215) 568-5549

TELEX: None

INFORMATION FOR SEQ ID NO: 48:

SEQUENCE CHARACTERISTICS:

LENGTH: 7 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-462-860-48

Query Match

Best Local Similarity 100.0%; Score 18; DB 8; Length 7;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
||||  
Db 2 SLDA 5

RESULT 4  
US-08-462-860A-48

; Sequence 48, Application US/08462860A

; GENERAL INFORMATION:

APPLICANT: Cooperman, Barry S.

APPLICANT: Rubin, Harvey

APPLICANT: Salem, Jerome

TITLE OF INVENTION: PEPTIDE INHIBITORS OF PLASMODIUM FALCIPARUM

TITLE OF INVENTION: RIBONUCLEOTIDE

NUMBER OF SEQUENCES: 69

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEIDEL, GONDA, LAVORIGNA & MONACO, P.C.

STREET: Suite 1800, Two Penn Center Plaza

CITY: Philadelphia

STATE: PA

COUNTRY: U.S.A.

ZIP: 19102-1786

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/462,860A

FILING DATE: 05-JUN-1995

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/136,743

FILING DATE: 14-OCT-1993

ATTORNEY/AGENT INFORMATION:

NAME: Monaco, Daniel A.

```
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 3957-10 D11
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-462-860A-48

Query Match
Best Local Similarity 100.0%; Score 18; DB 8; Length 7;
Pred. NO. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4
   ||||
Db 2 SLDA 5

RESULT 5
US-09-194-218-41
; SEQUENCE 41, Application US/09194218
; GENERAL INFORMATION:
; APPLICANT: KINGSTON, David J.
; APPLICANT: GERRATY, Norman L.
; APPLICANT: WESTBROOK, Simon L.
; TITLE OF INVENTION: - PEPTIDES, ANTIBODIES, VACCINES AND USES THEREOF
; FILE REFERENCE: 016786/0214
; CURRENT APPLICATION NUMBER: US/09/194,218
; EARLIER FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: PCT/AU97/00312
; EARLIER FILING DATE: 1997-05-22
; EARLIER APPLICATION NUMBER: AU P99990
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-194-218-41

Query Match
Best Local Similarity 100.0%; Score 18; DB 15; Length 7;
Pred. NO. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4
   ||||
Db 2 SLDA 5

RESULT 6
US-09-194-218-44
; SEQUENCE 44, Application US/09194218
; GENERAL INFORMATION:
; APPLICANT: KINGSTON, David J.
; APPLICANT: GERRATY, Norman L.
; APPLICANT: WESTBROOK, Simon L.
; TITLE OF INVENTION: MODULATING THE ACTIVITY OF HORMONES OR THEIR RECEPTORS
; FILE REFERENCE: 016786/0214
; CURRENT APPLICATION NUMBER: US/09/194,218
; EARLIER FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: PCT/AU97/00312
; EARLIER FILING DATE: 1997-05-22
; EARLIER APPLICATION NUMBER: AU P99990
; EARLIER FILING DATE: 1996-05-22
; NUMBER OF SEQ ID NOS: 58
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Rat
; US-09-194-218-44

Query Match
Best Local Similarity 100.0%; Score 18; DB 15; Length 7;
Pred. NO. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4
   ||||
Db 2 SLDA 5

RESULT 7
US-09-396-347B-53
; SEQUENCE 53, Application US/09396347B
; GENERAL INFORMATION:
; APPLICANT: Laal, Suman
; APPLICANT: Zolla-Pazner, Susan
; APPLICANT: Belisle, John T
; TITLE OF INVENTION: EARLY DETECTION OF MYCOBACTERIAL DISEASE
; FILE REFERENCE: NYU-11/CIP
; CURRENT APPLICATION NUMBER: US/09/396,347B
; EARLIER FILING DATE: 1999-09-14
; PRIOR APPLICATION NUMBER: 09/001,984
; PRIOR FILING DATE: 1997-12-31
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis strain H37Rv
; US-09-396-347B-53

Query Match
Best Local Similarity 100.0%; Score 18; DB 17; Length 7;
Pred. NO. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4
   ||||
Db 4 SLDA 7

RESULT 8
US-09-396-347D-53
; SEQUENCE 53, Application US/09396347D
; GENERAL INFORMATION:
; APPLICANT: Laal, Suman
; APPLICANT: Zolla-Pazner, Susan
; APPLICANT: Belisle, John T
; TITLE OF INVENTION: EARLY DETECTION OF MYCOBACTERIAL DISEASE
; FILE REFERENCE: 32004-169276
; CURRENT APPLICATION NUMBER: US/09/396,347D
; EARLIER FILING DATE: 1999-09-14
; PRIOR APPLICATION NUMBER: 09/001,984
; PRIOR FILING DATE: 1997-12-31
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis strain H37Rv
; US-09-396-347D-53

Query Match
Best Local Similarity 100.0%; Score 18; DB 17; Length 7;
Pred. NO. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4
   ||||
```

Db 4 SLDA 7

RESULT 9

US-09-396-347E-53

Sequence 53, Application US/09396347E

GENERAL INFORMATION:

APPLICANT: Laal, Suman

APPLICANT: Zolla-Pazner, Susan

APPLICANT: Belisle, John T

TITLE OF INVENTION: EARLY DETECTION OF MYCOBACTERIAL DISEASE

FILE REFERENCE: 32004-169276

CURRENT APPLICATION NUMBER: US/09/396,347E

PRIOR FILING DATE: 1999-09-14

PRIOR APPLICATION NUMBER: 09/001,984

NUMBER OF SEQ ID NOS: 106

SOFTWARE: Patentln Ver. 2.1

SEQ ID NO 53

LENGTH: 7

TYPE: PRT

ORGANISM: Mycobacterium tuberculosis strain H37Rv

US-09-396-347E-53

Query Match

Best Local Similarity 100.0%; Score 18; DB 17; Length 7;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4

Db 4 SLDA 7

RESULT 10

US-09-396-347F-53

Sequence 53, Application US/09396347F

GENERAL INFORMATION:

APPLICANT: Laal, Suman

APPLICANT: Zolla-Pazner, Susan

APPLICANT: Belisle, John T

TITLE OF INVENTION: EARLY DETECTION OF MYCOBACTERIAL DISEASE

FILE REFERENCE: 32004-169276

CURRENT APPLICATION NUMBER: US/09/396,347F

PRIOR FILING DATE: 1999-09-14

PRIOR APPLICATION NUMBER: 09/001,984

PRIOR FILING DATE: 1997-12-31

NUMBER OF SEQ ID NOS: 106

SOFTWARE: Patentln Ver. 2.1

SEQ ID NO 53

LENGTH: 7

TYPE: PRT

ORGANISM: Mycobacterium tuberculosis strain H37Rv

US-09-396-347F-53

Query Match

Best Local Similarity 100.0%; Score 18; DB 17; Length 7;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4

Db 4 SLDA 7

RESULT 11

US-09-758-128-41

Sequence 41, Application US/09758128

GENERAL INFORMATION:

APPLICANT: KINGSTON, David J.

APPLICANT: GERRATY, Norman L.

APPLICANT: WESTBROOK, Simon L.

TITLE OF INVENTION: MODULATING THE ACTIVITY OF HORMONES OR THEIR RECEPTORS

FILE REFERENCE: 016786/0214

PEPTIDES, ANTIBODIES, VACCINES AND USES THEREOF

CURRENT APPLICATION NUMBER: US/09/758,128

CURRENT FILING DATE: 2001-01-12

PRIOR APPLICATION NUMBER: 09/194,218

PRIOR FILING DATE: 1999-02-05

PRIOR APPLICATION NUMBER: AU PN9990

PRIOR FILING DATE: 1996-05-22

NUMBER OF SEQ ID NOS: 58

SOFTWARE: Patentln Ver. 2.0

SEQ ID NO 41

LENGTH: 7

TYPE: PRT

ORGANISM: Homo sapiens

US-09-758-128-41

Query Match

Best Local Similarity 100.0%; Score 18; DB 21; Length 7;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4

Db 2 SLDA 5

RESULT 12

US-09-758-128-44

Sequence 44, Application US/09758128

GENERAL INFORMATION:

APPLICANT: KINGSTON, David J.

APPLICANT: GERRATY, Norman L.

APPLICANT: WESTBROOK, Simon L.

TITLE OF INVENTION: MODULATING THE ACTIVITY OF HORMONES OR THEIR RECEPTORS

FILE REFERENCE: 016786/0214

CURRENT APPLICATION NUMBER: US/09/758,128

CURRENT FILING DATE: 2001-01-12

PRIOR APPLICATION NUMBER: 09/194,218

PRIOR FILING DATE: 1999-02-05

PRIOR APPLICATION NUMBER: AU PN9990

PRIOR FILING DATE: 1996-05-22

NUMBER OF SEQ ID NOS: 58

SOFTWARE: Patentln Ver. 2.0

SEQ ID NO 44

LENGTH: 7

TYPE: PRT

ORGANISM: Rat

US-09-758-128-44

Query Match

Best Local Similarity 100.0%; Score 18; DB 21; Length 7;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4

Db 2 SLDA 5

RESULT 13

US-09-758-198-41

Sequence 41, Application US/09758198

GENERAL INFORMATION:

APPLICANT: KINGSTON, David J.

APPLICANT: GERRATY, Norman L.

APPLICANT: WESTBROOK, Simon L.

TITLE OF INVENTION: MODULATING THE ACTIVITY OF HORMONES OR THEIR RECEPTORS

FILE REFERENCE: 016786/0214

CURRENT APPLICATION NUMBER: US/09/758,198

CURRENT FILING DATE: 2001-01-12

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/194,218

PRIOR FILING DATE: EARLIER FILING DATE: 1999-02-05

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PCT/AU97/00312

PRIOR FILING DATE: EARLIER FILING DATE: 1997-05-22

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: AU PN9990

PRIOR FILING DATE: EARLIER FILING DATE: 1996-05-22  
NUMBER OF SEQ ID NOS: 58  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 41  
LENGTH: 7  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-758-198-41

Query Match 100.0%; Score 18; DB 21; Length 7;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|||||  
DB 2 SLDA 5

RESULT 14  
US-09-758-198-44  
Sequence 44, Application US/09758198  
GENERAL INFORMATION:  
APPLICANT: KINGSTON, David J.  
APPLICANT: GERRATY, Norman L.  
APPLICANT: WESTBROOK, Simon L.  
TITLE OF INVENTION: - PEPTIDES, ANTIBODIES, VACCINES AND USES THEREOF  
FILE REFERENCE: 016786/0214  
CURRENT APPLICATION NUMBER: US/09/758,198  
PRIOR FILING DATE: 2001-01-12  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/194,218  
PRIOR FILING DATE: EARLIER FILING DATE: 1999-02-05  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PCT/AU97/00312  
PRIOR FILING DATE: EARLIER FILING DATE: 1997-05-22  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: AU PN9990  
NUMBER OF SEQ ID NOS: 58  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 44  
LENGTH: 7  
TYPE: PRT  
ORGANISM: Rat  
US-09-758-198-44

Query Match 100.0%; Score 18; DB 21; Length 7;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|||||  
DB 2 SLDA 5

RESULT 15  
US-09-758-426-41  
Sequence 41, Application US/09758426  
GENERAL INFORMATION:  
APPLICANT: KINGSTON, David J.  
APPLICANT: GERRATY, Norman L.  
APPLICANT: WESTBROOK, Simon L.  
TITLE OF INVENTION: MODULATING THE ACTIVITY OF HORMONES OR THEIR RECEPTORS  
FILE REFERENCE: 016786/0214  
CURRENT APPLICATION NUMBER: US/09/758,426  
PRIOR FILING DATE: 2001-01-12  
PRIOR APPLICATION NUMBER: 09/194,218  
PRIOR FILING DATE: 1999-02-05  
PRIOR APPLICATION NUMBER: AU PN9990  
PRIOR FILING DATE: 1996-05-22  
NUMBER OF SEQ ID NOS: 58  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 41  
LENGTH: 7

TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-758-426-41

Query Match 100.0%; Score 18; DB 21; Length 7;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|||||  
DB 2 SLDA 5

RESULT 16  
US-09-758-426-44  
Sequence 44, Application US/09758426  
GENERAL INFORMATION:  
APPLICANT: KINGSTON, David J.  
APPLICANT: GERRATY, Norman L.  
APPLICANT: WESTBROOK, Simon L.  
TITLE OF INVENTION: - PEPTIDES, ANTIBODIES, VACCINES AND USES THEREOF  
FILE REFERENCE: 016786/0214  
CURRENT APPLICATION NUMBER: US/09/758,426  
PRIOR FILING DATE: 2001-01-12  
PRIOR APPLICATION NUMBER: 09/194,218  
PRIOR FILING DATE: 1999-02-05  
PRIOR APPLICATION NUMBER: AU PN9990  
NUMBER OF SEQ ID NOS: 58  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 44  
LENGTH: 7  
TYPE: PRT  
ORGANISM: Rat  
US-09-758-426-44

Query Match 100.0%; Score 18; DB 21; Length 7;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|||||  
DB 2 SLDA 5

RESULT 17  
US-09-861-661-41  
Sequence 41, Application US/09861661  
GENERAL INFORMATION:  
APPLICANT: KINGSTON, DAVID J.  
APPLICANT: GERRATY, NORMAN L.  
APPLICANT: WESTBROOK, SIMON L.  
TITLE OF INVENTION: PEPTIDES, ANTIBODIES, VACCINES & USES THEREOF  
FILE REFERENCE: 054270/0135  
CURRENT APPLICATION NUMBER: US/09/861,661  
PRIOR FILING DATE: 2001-05-22  
PRIOR APPLICATION NUMBER: 09/194,218  
PRIOR FILING DATE: 1999-02-05  
PRIOR APPLICATION NUMBER: AU PN9990  
PRIOR FILING DATE: 1996-05-22  
NUMBER OF SEQ ID NOS: 59  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 41  
LENGTH: 7  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-861-661-41

Query Match 100.0%; Score 18; DB 22; Length 7;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
||||  
Db 2 SLDA 5

RESULT 18  
US-09-861-661-44

; Sequence 44, Application US/09861661  
; GENERAL INFORMATION:  
; APPLICANT: KINGSTON, DAVID J.  
; APPLICANT: GERRATY, NORMAN L.  
; APPLICANT: WESTBROOK, SIMON L.  
; TITLE OF INVENTION: PEPTIDES, ANTIBODIES, VACCINES & USES THEREOF  
; FILE REFERENCE: 054270/0135  
; CURRENT APPLICATION NUMBER: US/09/861,661  
; PRIOR FILING DATE: 2001-05-22  
; PRIOR APPLICATION NUMBER: 09/194,218  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: AU/PN9990  
; PRIOR FILING DATE: 1996-05-22  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 44  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Rattus sp.  
US-09-861-661-44

Query Match 100.0%; Score 18; DB 22; Length 7;  
Best Local Similarity 100.0%; Pred. NO. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
||||  
Db 2 SLDA 5

RESULT 19  
US-09-997-900-10

; Sequence 10, Application US/09997900  
; GENERAL INFORMATION:  
; APPLICANT: Kakefuda, Genichi  
; APPLICANT: Costello, Colleen  
; APPLICANT: Sun, Ming  
; APPLICANT: Hu, Weiming  
; TITLE OF INVENTION: Genes and Vectors for Confering Herbicide Resistance  
; FILE REFERENCE: 043753/241148 (5849-20A)  
; CURRENT APPLICATION NUMBER: US/09/997,900  
; PRIOR FILING DATE: 2001-11-30  
; PRIOR APPLICATION NUMBER: 60/106,239  
; PRIOR FILING DATE: 1998-10-29  
; PRIOR APPLICATION NUMBER: 09/426,568  
; PRIOR FILING DATE: 1999-10-22  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Arabidopsis sp.  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(7)  
; OTHER INFORMATION: N-terminal sequence of AHAS small subunit peptide  
; OTHER INFORMATION: of plasmid F2  
; NAME/KEY: SITE  
; LOCATION: (1)..(2)  
; OTHER INFORMATION: Thrombin cleavage site  
US-09-997-900-10

Query Match 100.0%; Score 18; DB 23; Length 7;  
Best Local Similarity 100.0%; Pred. NO. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
||||  
Db 2 SLDA 5

RESULT 20  
US-08-526-708-12

; Sequence 12, Application US/08526708  
; GENERAL INFORMATION:  
; APPLICANT: Ruostaheti, Erkki  
; APPLICANT: Pasqualini, Renata  
; TITLE OF INVENTION: Molecules That Home to a Selected Organ  
; TITLE OF INVENTION: In Vivo  
; NUMBER OF SEQUENCES: 46  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell & Flores LLP  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: United States  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/526,708  
; FILING DATE: 11-SEP-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LJ 1802  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8 amino acids  
; TYPE: amino acid  
; TOPOLOGY: both  
; MOLECULE TYPE: peptide  
US-08-526-708-12

Query Match 100.0%; Score 18; DB 9; Length 8;  
Best Local Similarity 100.0%; Pred. NO. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
||||  
Db 4 SLDA 7

RESULT 21  
US-08-813-273-12

; Sequence 12, Application US/08813273  
; GENERAL INFORMATION:  
; APPLICANT: Ruostaheti, Erkki  
; APPLICANT: Pasqualini, Renata  
; TITLE OF INVENTION: Method of Identifying Molecules That  
; TITLE OF INVENTION: Home to a Selected Organ or Tissue In Vivo  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell & Flores LLP  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: United States  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/813,273  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 2410  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ. ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-813-273-12

Query Match 100.0%; Score 18; DB 12; Length 8;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
1111  
DB 4 SLDA 7

RESULT 22  
US-09-228-866-12  
Sequence 12, Application US/09228866  
GENERAL INFORMATION:  
APPLICANT: Ruoslahti, Erkki  
TITLE OF INVENTION: Molecules That Home to a Selected Organ  
NUMBER OF INVENTION: In Vivo  
NUMBER OF SEQUENCES: 46  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell & Flores LLP  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: United States  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/228,866  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/526,708  
FILING DATE: 11-SEP-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 3430  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ. ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
TOPOLOGY: both  
MOLECULE TYPE: peptide  
US-09-228-866-12

Query Match 100.0%; Score 18; DB 16; Length 8;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
1111  
DB 4 SLDA 7

RESULT 23  
US-09-700-993-25  
Sequence 25, Application US/09700993  
GENERAL INFORMATION:  
APPLICANT: KOBAYASHI, KAZUO  
APPLICANT: TAKEUCHI, MAKOTO  
APPLICANT: IWAMATSU, AKIHIKO  
APPLICANT: YAMAMOTO, KENJI  
APPLICANT: KUMAGAI, HIDEHIKO  
APPLICANT: YOSHIDA, SATOSHI  
TITLE OF INVENTION: ENDO-BETA-N-ACETYLGUCOSAMINIDASE GENE  
FILE REFERENCE: 081356/0154  
CURRENT APPLICATION NUMBER: US/09/700,993  
CURRENT FILING DATE: 2000-11-22  
PRIOR APPLICATION NUMBER: JP98/141717  
PRIOR FILING DATE: 1998-05-22  
NUMBER OF SEQ ID NOS: 37  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 25  
LENGTH: 8  
TYPE: PRT  
ORGANISM: Mucor hiemalis  
US-09-700-993-25

Query Match 100.0%; Score 18; DB 21; Length 8;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
1111  
DB 3 SLDA 6

RESULT 24  
US-09-922-227-12  
Sequence 12, Application US/09922227  
GENERAL INFORMATION:  
APPLICANT: Ruoslahti, Erkki  
TITLE OF INVENTION: Method of Identifying Molecules That Home to a Selected Organ In Vivo  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell & Flores LLP  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: United States  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/922,227  
FILING DATE: 02-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/526,710  
FILING DATE: 11-SEP-1995  
APPLICATION NUMBER: US 08/813,273  
FILING DATE: 10-MAR-1997

APPLICATION NUMBER: US 08/862,855  
FILING DATE: 23-MAY-1997  
APPLICATION NUMBER: US 09/227,906  
FILING DATE: 08-JAN-1999  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 4859  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-09-922-227-12

Query Match 100.0%; Score 18; DB 23; Length 8;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
Db 4 SLDA 7

RESULT 25  
US-09-988-493-271  
Sequence 271, Application US/09988493  
GENERAL INFORMATION:  
APPLICANT: Herath, Herath Mudiyanselage Athula Chandrasiri  
APPLICANT: O'Hare, Michael John  
APPLICANT: Page, Martin John  
APPLICANT: Parekh, Rajesh Bhikhu  
APPLICANT: Waterfield, Michael Derek  
TITLE OF INVENTION: Proteins, genes, and Their Use for  
FILE OF INVENTION: Diagnosis and Treatment of Breast Cancer  
FILE REFERENCE: 2543-1-024  
CURRENT APPLICATION NUMBER: US/09/988,493  
CURRENT FILING DATE: 2002-05-21  
PRIOR APPLICATION NUMBER: PCT/GB01/01219  
PRIOR FILING DATE: 2001-03-20  
PRIOR APPLICATION NUMBER: GB 0006695.1  
PRIOR FILING DATE: 2000-03-20  
PRIOR APPLICATION NUMBER: GB 0007265.2  
PRIOR FILING DATE: 2000-02-24  
NUMBER OF SEQ ID NOS: 308  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 271  
LENGTH: 8  
TYPE: PRT  
ORGANISM: homo sapien  
US-09-988-493-271

Query Match 100.0%; Score 18; DB 23; Length 8;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
Db 1 SLDA 4

RESULT 26  
US-10-014-340-647  
Sequence 647, Application US/10014340  
GENERAL INFORMATION:  
APPLICANT: Herath, et al  
TITLE OF INVENTION: Nucleic Acid Molecules, Polypeptides and Uses Therefor, Including  
TITLE OF INVENTION: Diagnosis and Treatment of Alzheimer's Disease

FILE REFERENCE: 9195-078  
CURRENT APPLICATION NUMBER: US/10/014,340  
CURRENT FILING DATE: 2001-12-10  
NUMBER OF SEQ ID NOS: 823  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 647  
LENGTH: 8  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-014-340-647

Query Match 100.0%; Score 18; DB 24; Length 8;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
Db 1 SLDA 4

RESULT 27  
PCT-US97-21437-13  
Sequence 13, Application PC/TUS9721437,  
GENERAL INFORMATION:  
APPLICANT: The General Hospital Corporation  
TITLE OF INVENTION: MODULATION OF IGC BINDING  
TITLE OF INVENTION: TO FcRn  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US97/21437  
FILING DATE: 26-NOV-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/031,607  
FILING DATE: 27-NOV-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 34,819  
REFERENCE/DOCKET NUMBER: 00786/360W02  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-542-5070  
TELEFAX: 617-542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
PCT-US97-21437-13

Query Match 100.0%; Score 18; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
Db 7 SLDA 10

RESULT 28  
US-09-109-836-13  
Sequence 13, Application US/09109836



GENERAL INFORMATION:  
APPLICANT: Jacobowitz Israel, Esther  
APPLICANT: Simister, Neil E.  
TITLE OF INVENTION: MODULATION OF IGG BINDING  
TITLE OF INVENTION: TO FCRI  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/109,836  
FILING DATE: 02-JUL-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/977,836  
FILING DATE: 26-NOV-1997  
APPLICATION NUMBER: 60/031,607  
FILING DATE: 27-NOV-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 34,819  
REFERENCE/DOCKET NUMBER: 00786/360003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-542-5070  
TELEFAX: 617-542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-09-109-836-13

Query Match  
Best Local Similarity 100.0%; Score 18; DB 15; Length 10;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
|||||  
Db 7 SLDA 10

RESULT 29  
US-08-188-354-2  
Sequence 2, Application US/08188354  
GENERAL INFORMATION:  
APPLICANT: COMAN, Nicholas J.  
TITLE OF INVENTION: NOVEL EUKARYOTIC FOLDING PROTEINS, NUCLEIC  
ACID SEQUENCES, AND METHODS OF MAKING AND USING THEREOF  
ENCODING THEREFOR  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEIMARK  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/188,354  
FILING DATE: 28-JAN-1994

CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/897,162  
FILING DATE: 11-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/05568  
FILING DATE: 11-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: TOWNSEND, G. Kevin  
REGISTRATION NUMBER: 34,033  
REFERENCE/DOCKET NUMBER: COMAN=3A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-188-354-2

Query Match  
Best Local Similarity 100.0%; Score 18; DB 5; Length 12;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
|||||  
Db 6 SLDA 9

RESULT 30  
US-09-831-253A-3  
Sequence 3, Application US/09831253A  
GENERAL INFORMATION:  
APPLICANT: EZQUERRO SAENZ, Ignacio Jose  
APPLICANT: LASARTE SAGASTIBELZA, Juan Jose  
APPLICANT: PRIETO VALTUENA, Jesus  
APPLICANT: BORRAS CUESTA, Francisco  
TITLE OF INVENTION: TGF-beta1-inhibitor peptides  
FILE REFERENCE: U013446-9  
CURRENT APPLICATION NUMBER: US/09/831,253A  
FILING DATE: 2001-05-07  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3  
LENGTH: 12  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: TGF-beta1,  
US-09-831-253A-3

Query Match  
Best Local Similarity 100.0%; Score 18; DB 22; Length 12;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
|||||  
Db 2 SLDA 5

RESULT 31  
US-08-461-383-96  
Sequence 96, Application US/08461383  
GENERAL INFORMATION:  
APPLICANT: FOWLES, Dana M.  
APPLICANT: BROACH, Jim  
APPLICANT: MANFREDI, John  
APPLICANT: KLEIN, Christine

APPLICANT: MURPHY, Andrew J.  
APPLICANT: PAUL, Jeremy  
APPLICANT: TRUEHEART, Joshua  
TITLE OF INVENTION: YEAST CELLS ENGINEERED TO PRODUCE  
TITLE OF INVENTION: PHERMONE SYSTEM PROTEIN SURROGATES, AND USES THEREFOR  
NUMBER OF SEQUENCES: 119  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE AND COCKFIELD  
STREET: 60 State Street, Suite 510  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/461,383  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/322,137  
FILING DATE: 13-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/309,313  
FILING DATE: 20-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/190,328  
FILING DATE: 31-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/041,431  
FILING DATE: 31-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: Cpl-012CP4B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-227-7400  
TELEFAX: 617-227-5941  
TELEX: 752806  
INFORMATION FOR SEQ ID NO: 96:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-461-383-96  
Query Match  
Best Local Similarity 100.0%; Score 18; DB 8; Length 13;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SLDA 4  
Db 5 SLDA 8  
RESULT 32  
US-08-463-181-96  
Sequence 96, Application US/08463181  
GENERAL INFORMATION:  
APPLICANT: FOWLKES, Dana M.  
APPLICANT: BROACH, Jim  
APPLICANT: MANFREDI, John  
APPLICANT: KLEIN, Christine  
APPLICANT: MURPHY, Andrew J.  
APPLICANT: PAUL, Jeremy  
APPLICANT: TRUEHEART, Joshua  
TITLE OF INVENTION: YEAST CELLS ENGINEERED TO PRODUCE  
TITLE OF INVENTION: PHERMONE SYSTEM PROTEIN SURROGATES, AND USES THEREFOR  
NUMBER OF SEQUENCES: 119

CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEWMARK  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/463,181  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/322,137  
FILING DATE: 13-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/309,313  
FILING DATE: 20-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/190,328  
FILING DATE: 31-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/041,431  
FILING DATE: 31-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: COOPER, Iver P.  
REGISTRATION NUMBER: 28,005  
REFERENCE/DOCKET NUMBER: FOLKRES-2D  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 96:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-463-181-96  
Query Match  
Best Local Similarity 100.0%; Score 18; DB 8; Length 13;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SLDA 4  
Db 5 SLDA 8  
RESULT 33  
US-08-689-172A-32  
Sequence 32, Application US/08689172A  
GENERAL INFORMATION:  
APPLICANT: Klein, Christine A.  
APPLICANT: Murphy, Andrew J. M.  
TITLE OF INVENTION: Methods and Compositions for  
TITLE OF INVENTION: Identifying Receptor Effectors  
NUMBER OF SEQUENCES: 112  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street, Suite 510  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Ascii(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/689,172A  
FILING DATE: 07-Aug-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: KARA, Catherine J.  
REGISTRATION NUMBER: P41,106  
REFERENCE/DOCKET NUMBER: CPI-012CP7  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-689-172A-32

Query Match 100.0%; Score 18; DB 10; Length 13;  
Best Local Similarity 100.0%; Pred. No. 6.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
||||  
DB 5 SLDA 8

RESULT 34  
US-08-689-172C-32  
Sequence 32, Application US/08689172C  
GENERAL INFORMATION:  
APPLICANT: Klein, Christine A.  
APPLICANT: Murphy, Andrew J. M.  
TITLE OF INVENTION: Methods and Compositions for  
IDENTIFYING RECEPTOR EFFECTORS  
NUMBER OF SEQUENCES: 132  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Ascii(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/689,172C  
FILING DATE: 07-Aug-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: KARA, Catherine J.  
REGISTRATION NUMBER: P41,106  
REFERENCE/DOCKET NUMBER: CPI-012CP7  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-689-172C-32

Query Match 100.0%; Score 18; DB 10; Length 13;  
Best Local Similarity 100.0%; Pred. No. 6.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
||||  
DB 5 SLDA 8

RESULT 35  
US-09-286-166-96  
Sequence 96, Application US/09286166  
GENERAL INFORMATION:  
APPLICANT: FOWLES, Dana M.  
APPLICANT: BROACH, Jim  
APPLICANT: MANFREDI, John  
APPLICANT: KLEIN, Christine  
APPLICANT: MURPHY, Andrew J.  
APPLICANT: PAUL, Jeremy  
APPLICANT: TRUEHEART, Joshua  
TITLE OF INVENTION: YEAST CELLS ENGINEERED TO PRODUCE  
PHERMONE SYSTEM PROTEIN SURROGATES, AND USES THEREFOR  
NUMBER OF SEQUENCES: 119  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE AND COCKFIELD  
STREET: 60 State Street, Suite 510  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/286,166  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/461,383  
FILING DATE: 05-JUN-1995  
APPLICATION NUMBER: US 08/322,137  
FILING DATE: 13-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/309,313  
FILING DATE: 20-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/190,328  
FILING DATE: 31-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/041,431  
FILING DATE: 31-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: CPI-012CP4B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-227-7400  
TELEFAX: 617-227-5941  
TELEX: 752806  
INFORMATION FOR SEQ ID NO: 96:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-286-166-96

Query Match 100.0%; Score 18; DB 16; Length 13;  
Best Local Similarity 100.0%; Pred. No. 6.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
||||  
DB 5 SLDA 8

```
RESULT 36
US-09-309-196-96
; Sequence 96, Application US/09309196
; GENERAL INFORMATION:
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: BROACH, Jim
; APPLICANT: MANFREDI, John
; APPLICANT: KLEIN, Christine
; APPLICANT: MURPHY, Andrew J.
; APPLICANT: PAUL, Jeremy
; APPLICANT: TRUEHEART, Joshua
; TITLE OF INVENTION: YEAST CELLS ENGINEERED TO PRODUCE
; PERMOE SYSTEM PROTEIN SURROGATES, AND USES THEREFOR
; NUMBER OF SEQUENCES: 119
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/309,196
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/322,137
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/190,328
; FILING DATE: 31-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/041,431
; FILING DATE: 31-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: COOPER, Ivet P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: FOLWRKS-2C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 96:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-309-196-96

Query Match      100.0%; Score 18; DB 17; Length 13;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4
   ||||
Db 5 SLDA 8

RESULT 37
US-09-747-774A-32
; Sequence 32, Application US/09747774A
; GENERAL INFORMATION:
; APPLICANT: Klein, Christine A.
; APPLICANT: Murphy, Andrew J. M.
; APPLICANT: Broach, James R.
; APPLICANT: Manfredi, John
```

```
; APPLICANT: Paul, Jeremy
; APPLICANT: Fowlkes, Dana M.
; APPLICANT: Trueheart, Joshua
; TITLE OF INVENTION: Methods and Compositions for Identifying
; FILE REFERENCE: CPI-012CP5DV
; CURRENT APPLICATION NUMBER: US/09/747,774A
; PRIOR APPLICATION NUMBER: US 08/582,333
; PRIOR FILING DATE: 1996-01-17
; PRIOR APPLICATION NUMBER: US 08/464,531
; PRIOR FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: US 08/461,598
; PRIOR FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: US 08/461,383
; PRIOR FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: US 08/463,181
; PRIOR FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: US 08/322,137
; PRIOR FILING DATE: 1994-10-13
; PRIOR APPLICATION NUMBER: US 08/309,313
; PRIOR FILING DATE: 1994-09-20
; PRIOR APPLICATION NUMBER: US 08/190,328
; PRIOR FILING DATE: 1994-01-31
; PRIOR APPLICATION NUMBER: US 08/041,431
; PRIOR FILING DATE: 1993-03-31
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-747-774A-32

Query Match      100.0%; Score 18; DB 21; Length 13;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4
   ||||
Db 5 SLDA 8

RESULT 38
US-09-953-354-32
; Sequence 32, Application US/09953354
; GENERAL INFORMATION:
; APPLICANT: Klein, Christine A.
; APPLICANT: Murphy, Andrew J. M.
; TITLE OF INVENTION: Methods and Compositions for
; Identifying Receptor Effectors
; NUMBER OF SEQUENCES: 132
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Ascii(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/953,354
; FILING DATE: 13-Sep-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/689,172
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: KARA, Catherine J.
```

REGISTRATION NUMBER: P41,106  
REFERENCE/DOCKET NUMBER: CRI-012CE7  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 32:  
US-09-953-354-32

Query Match  
Best Local Similarity 100.0%; Score 18; DB 23; Length 13;  
Pred. No. 6.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
|||||  
DB 5 SLDA 8

RESULT 39  
US-10-263-341-96  
Sequence 96, Application US/10263341  
GENERAL INFORMATION:  
APPLICANT: FOWLES, Dana M.  
BROACH, Jim  
MANFREDI, John  
KLEIN, Christine  
MURPHY, Andrew J.  
PAUL, Jeremy  
TRUEHEART, Joshua  
TITLE OF INVENTION: YEAST CELLS ENGINEERED TO PRODUCE  
PHEROMONE SYSTEM PROTEIN SURROGATES, AND USES THEREFOR  
NUMBER OF SEQUENCES: 119  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEIMARK  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/263,341  
FILING DATE: 01-Oct-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/322,137  
FILING DATE: 13-Oct-1994  
APPLICATION NUMBER: US 08/309,313  
FILING DATE: 20-Sep-1994  
APPLICATION NUMBER: US 08/190,328  
FILING DATE: 31-Jan-1994  
APPLICATION NUMBER: US 08/041,431  
FILING DATE: 31-Mar-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: COOPER, Iver P.  
REGISTRATION NUMBER: 28,005  
REFERENCE/DOCKET NUMBER: FOLWKS-2C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 96:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 96:  
US-10-263-341-96

Query Match  
Best Local Similarity 100.0%; Score 18; DB 26; Length 13;  
Pred. No. 6.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
|||||  
DB 5 SLDA 8

RESULT 40  
US-09-831-253A-6  
Sequence 6, Application US/09831253A  
GENERAL INFORMATION:  
APPLICANT: EZQUERRO SAENZ, Ignacio Jose  
APPLICANT: LASARTE SAGASTIBELZA, Juan Jose  
APPLICANT: PRIETO VALTUENA, Jesus  
APPLICANT: BORRAS CUESTA, Francisco  
TITLE OF INVENTION: TGF (b1)-inhibitor peptides  
FILE REFERENCE: U013446-9  
CURRENT APPLICATION NUMBER: US/09/831,253A  
CURRENT FILING DATE: 2001-05-07  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: Patent Ver. 2.0  
SEQ ID NO 6  
LENGTH: 14  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Derived from  
OTHER INFORMATION: the modified human type III receptor, position  
US-09-831-253A-6

Query Match  
Best Local Similarity 100.0%; Score 18; DB 22; Length 14;  
Pred. No. 7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
|||||  
DB 2 SLDA 5

RESULT 41  
US-08-400-306-1  
Sequence 1, Application US/08400306  
GENERAL INFORMATION:  
APPLICANT: J.A. Kessler, A.J. Conley, B.A. Arnold  
TITLE OF INVENTION: Selected Principal Neutralization  
TITLE OF INVENTION: Epitopes for Antibody that Neutralizes HIV in Peripheral  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merck & Co., Inc.  
STREET: P.O. Box 2000  
CITY: Rahway  
STATE: NJ  
COUNTRY: USA  
ZIP: 07065

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/400,306  
FILING DATE:  
CLASSIFICATION: 424

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/129,997  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Meredith, Roy D.  
REGISTRATION NUMBER: 30,777  
REFERENCE/DOCKET NUMBER: 19071  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 594-4678  
TELEFAX: (908) 594-4720  
TELEX: 138825  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
IMMEDIATE SOURCE: Random Epitope Library Alpha  
US-08-400-306-1

Query Match  
Best Local Similarity 100.0%; Score 18; DB 8; Length 15;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
1111  
Db 7 SLDA 10

RESULT 42  
PCT-US99-28465-15  
Sequence 15, Application PC/TUS9928465A  
GENERAL INFORMATION:  
APPLICANT: Temple University - Of The Commonwealth System of  
TITLE OF INVENTION: Inhibition of Angiogenesis By High Molecular Weight  
FILE REFERENCE: 6056-260 PC  
CURRENT APPLICATION NUMBER: PCT/US99/28465A  
CURRENT FILING DATE: 1999-12-02  
EARLIER APPLICATION NUMBER: 60/112,427  
EARLIER FILING DATE: 1998-12-16  
NUMBER OF SEQ ID NOS: 21  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 15  
LENGTH: 16  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Fragment of  
PCT-US99-28465-15

Query Match  
Best Local Similarity 100.0%; Score 18; DB 1; Length 16;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
1111  
Db 8 SLDA 11

RESULT 43  
US-09-461-061-15  
Sequence 15, Application US/09461061  
GENERAL INFORMATION:  
APPLICANT: McCrae, Keith R.  
TITLE OF INVENTION: Inhibition of Angiogenesis By High Molecular Weight  
FILE REFERENCE: 6056-260 US  
CURRENT APPLICATION NUMBER: US/09/461,061

CURRENT FILING DATE: 1999-12-15  
EARLIER APPLICATION NUMBER: 60/112,427  
EARLIER FILING DATE: 1998-12-16  
NUMBER OF SEQ ID NOS: 21  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 15  
LENGTH: 16  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Fragment of  
US-09-461-061-15

Query Match  
Best Local Similarity 100.0%; Score 18; DB 18; Length 16;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
1111  
Db 8 SLDA 11

RESULT 44  
US-09-461-061A-15  
Sequence 15, Application US/09461061A  
GENERAL INFORMATION:  
APPLICANT: McCrae, Keith R.  
TITLE OF INVENTION: Inhibition of Angiogenesis By High Molecular Weight  
FILE REFERENCE: 6056-260 US  
CURRENT APPLICATION NUMBER: US/09/461,061A  
CURRENT FILING DATE: 1999-12-15  
PRIOR APPLICATION NUMBER: 60/112,427  
PRIOR FILING DATE: 1998-12-16  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 15  
LENGTH: 16  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Fragment of  
US-09-461-061A-15

Query Match  
Best Local Similarity 100.0%; Score 18; DB 18; Length 16;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
1111  
Db 8 SLDA 11

RESULT 45  
PCT-US02-09671-65  
Sequence 65, Application PC/TUS0209671  
GENERAL INFORMATION:  
APPLICANT: Zygos Inc.  
TITLE OF INVENTION: TRANSLATIONAL PROFILING  
FILE REFERENCE: 08191-026W01  
CURRENT APPLICATION NUMBER: PCT/US02/09671  
CURRENT FILING DATE: 2002-03-28  
PRIOR APPLICATION NUMBER: 60/279,495  
PRIOR FILING DATE: 2001-03-28  
PRIOR APPLICATION NUMBER: 60/292,544  
PRIOR FILING DATE: 2001-05-21  
PRIOR APPLICATION NUMBER: 60/310,801  
PRIOR FILING DATE: 2001-08-08  
PRIOR APPLICATION NUMBER: 60/326,370  
PRIOR FILING DATE: 2001-10-01  
PRIOR APPLICATION NUMBER: 60/336,780

Thu Feb 6 12:30:15 2003

pat943-4.rapm

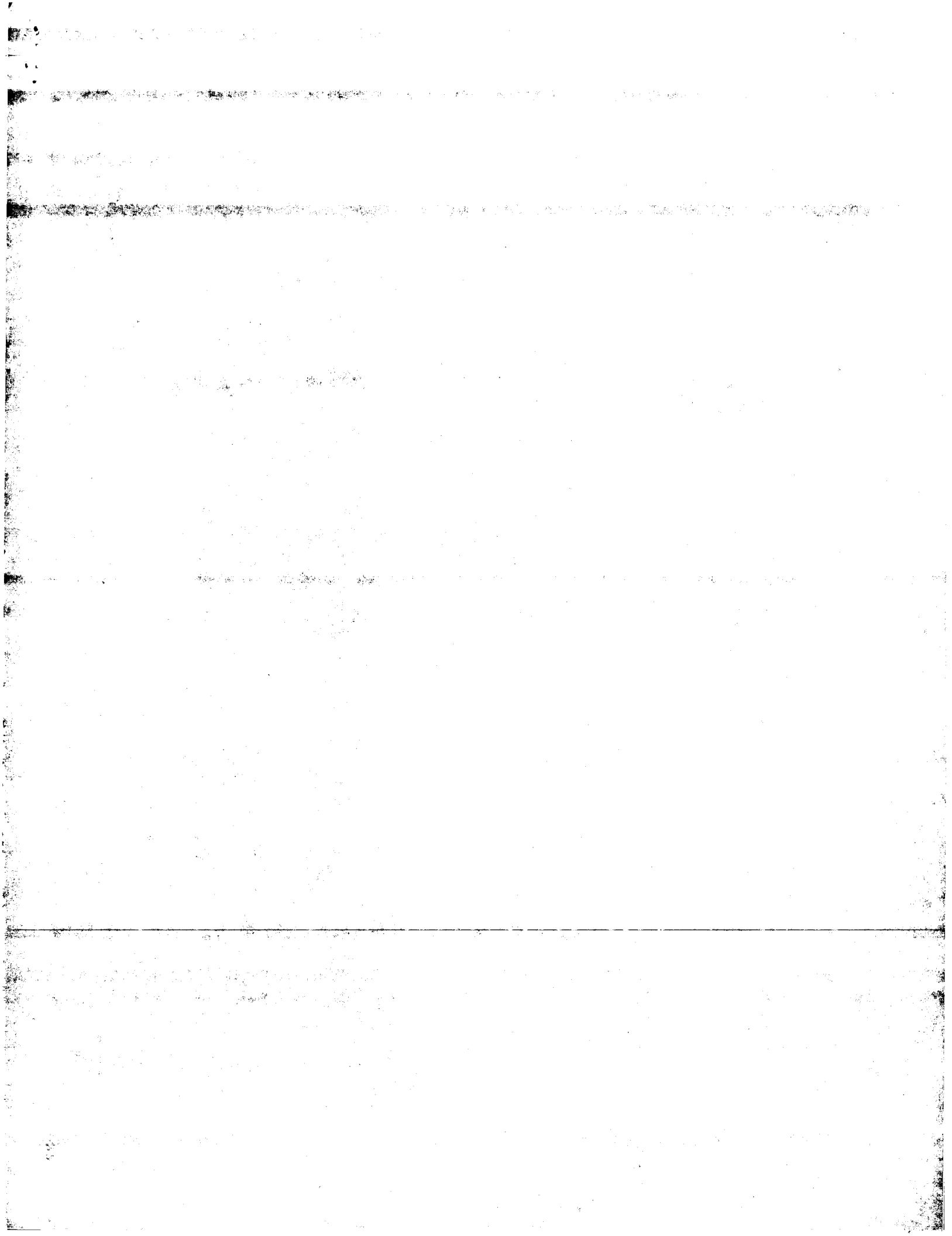
Page 15

; PRIOR FILING DATE: 2001-12-04  
; PRIOR APPLICATION NUMBER: 60/358,985  
; PRIOR FILING DATE: 2002-02-20  
; NUMBER OF SEQ ID NOS: 2041  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 65  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US02-09671-65

Query Match 100.0%; Score 18; DB 1; Length 17;  
Best Local Similarity 100.0%; Pred. No. 8.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
1111  
Db 7 SLDA 10

Search completed: February 6, 2003, 11:37:34  
Job time : 134.167 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: February 6, 2003, 11:19:49 ; Search time 14 Seconds  
(without alignments)  
23.165 Million cell updates/sec

Title: PAT943-4  
Perfect score: 18  
Sequence: 1 slide 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 405691 seqs, 81078759 residues

Total number of hits satisfying chosen parameters: 1588

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	100.0	13	US-09-258-600-96	Sequence 96, Appl
2	18	100.0	44	US-10-203-138A-12434	Sequence 12434, A
3	18	100.0	45	US-09-585-645A-54	Sequence 54, Appl
4	18	100.0	50	PCT-US02-32727-24214	Sequence 24214, A
5	18	100.0	50	US-10-057-498-24214	Sequence 24214, A
6	18	100.0	52	US-09-513-999C-5438	Sequence 5438, Ap
7	18	100.0	53	PCT-US02-32727-22975	Sequence 22975, A
8	18	100.0	53	US-10-057-498-22975	Sequence 22975, A
9	18	100.0	54	PCT-US02-32727-17389	Sequence 17389, A
10	18	100.0	54	US-10-057-498-17389	Sequence 17389, A
11	18	100.0	55	US-10-203-138A-11775	Sequence 11775, A
12	18	100.0	56	PCT-US02-32727-2538	Sequence 2538, Ap
13	18	100.0	56	PCT-US02-32727-27113	Sequence 27113, A
14	18	100.0	56	US-10-057-498-2538	Sequence 2538, Ap
15	18	100.0	56	US-10-057-498-27113	Sequence 27113, A
16	18	100.0	57	PCT-US02-32727-23045	Sequence 23045, A
17	18	100.0	57	PCT-US02-32727-27575	Sequence 27575, A
18	18	100.0	57	US-10-057-498-23045	Sequence 23045, A
19	18	100.0	57	US-10-057-498-27575	Sequence 27575, A
20	18	100.0	59	PCT-US02-32727-16333	Sequence 16333, A
21	18	100.0	59	PCT-US02-32727-28196	Sequence 28196, A
22	18	100.0	59	PCT-US02-32727-28948	Sequence 28948, A
23	18	100.0	59	US-10-057-498-16333	Sequence 16333, A
24	18	100.0	59	US-10-057-498-28196	Sequence 28196, A
25	18	100.0	59	US-10-057-498-28948	Sequence 28948, A
26	18	100.0	61	PCT-US02-32727-21639	Sequence 21639, A

27	18	100.0	61	US-10-057-498-21639	Sequence 21639, A
28	18	100.0	62	US/09/513	Sequence 7714, Ap
29	18	100.0	63	PCT-US02-32727-6416	Sequence 6416, Ap
30	18	100.0	63	US-10-301-997-94	Sequence 94, Appl
31	18	100.0	63	US-10-057-498-6416	Sequence 6416, Ap
32	18	100.0	65	PCT-US02-32727-8518	Sequence 8518, Ap
33	18	100.0	65	US-10-057-498-8518	Sequence 8518, Ap
34	18	100.0	68	PCT-US02-32727-2064	Sequence 2064, Ap
35	18	100.0	68	US-10-057-498-2064	Sequence 2064, Ap
36	18	100.0	70	PCT-US02-32727-6955	Sequence 6955, Ap
37	18	100.0	70	PCT-US02-32727-28376	Sequence 28376, A
38	18	100.0	70	US-10-057-498-6955	Sequence 6955, Ap
39	18	100.0	70	US-10-057-498-28376	Sequence 28376, A
40	18	100.0	72	US-10-203-138A-11100	Sequence 11100, A
41	18	100.0	72	PCT-US02-32727-13114	Sequence 13114, A
42	18	100.0	72	PCT-US02-32727-27093	Sequence 27093, A
43	18	100.0	72	US-10-264-213-210	Sequence 210, App
44	18	100.0	72	US-10-057-498-13114	Sequence 13114, A
45	18	100.0	72	US-10-057-498-27093	Sequence 27093, A

## ALIGNMENTS

RESULT 1  
US-09-258-600-96.  
Sequence 96, Application US/09258600  
GENERAL INFORMATION:  
APPLICANT: FOWLKES, Dana M.  
BROACH, Jim  
MANFREDI, John  
KLEIN, Christine  
MURPHY, Andrew J.  
PAUL, Jeremy  
TRUEHEART, Joshua  
TITLE OF INVENTION: YEAST CELLS ENGINEERED TO PRODUCE  
PHEROMONE SYSTEM PROTEIN SURROGATES, AND USES THEREFOR  
NUMBER OF SEQUENCES: 119  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: BROWDY AND NEIMARK  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/258, 600  
FILING DATE: 26-Feb-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/461, 598  
FILING DATE: 05-JUN-1995  
APPLICATION NUMBER: US 08/322, 137  
FILING DATE: 13-OCT-1994  
APPLICATION NUMBER: US 08/309, 313  
FILING DATE: 20-SEP-1994  
APPLICATION NUMBER: US 08/190, 328  
FILING DATE: 31-JAN-1994  
APPLICATION NUMBER: US 08/041, 431  
ATTORNEY/AGENT INFORMATION:  
NAME: COOPER, Iyer P.  
REGISTRATION NUMBER: 28, 005  
REFERENCE/DOCKET NUMBER: FOWLKES-2P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633

INFORMATION FOR SEQ ID NO: 96:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 96:  
US-09-258-600-96

Query Match 100.0%; Score 18; DB 5; Length 13;  
Best Local Similarity 100.0%; Pred. No. 74;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
||||  
DB 5 SLDA 8

RESULT 2  
US-10-203-138A-12434  
Sequence 12434, Application US/10203138A  
GENERAL INFORMATION:  
APPLICANT: Molecular Dynamics, Inc.  
APPLICANT: Penn. Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: PB 0004 WO 8  
CURRENT APPLICATION NUMBER: US/10/203,138A  
CURRENT FILING DATE: 2002-08-02  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 04 February 2000 (04.02.00)  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 26 May 2000 (26.05.00)  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 03 August 2000 (03.08.00)  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 03 October 2000 (03.10.00)  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 27 September 2000 (27.09.00)  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 21 September 2000 (21.09.00)  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 30 June 2000 (30.06.00)  
NUMBER OF SEQ ID NOS: 15438  
SOFTWARE: Molecular Dynamics Sequence Listing Engine  
SEQ ID NO 12434  
LENGTH: 44  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC005488.2  
FEATURE:  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.2  
FEATURE:  
OTHER INFORMATION: SWISSPROT HIT: P54278, EVALUATE 4.00e-16  
FEATURE:  
OTHER INFORMATION: EST\_HUMAN HIT: BF680924.1, EVALUATE 1.00e-17  
US-10-203-138A-12434

Query Match 100.0%; Score 18; DB 6; Length 44;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
||||  
DB 36 SLDA 39

RESULT 3  
US-09-585-645A-54

Sequence 54, Application US/09585645A  
GENERAL INFORMATION:  
APPLICANT: Zoghbi, Huda  
APPLICANT: Bellien, Hugo  
APPLICANT: Birmingham, Nessim  
APPLICANT: Hassan, Bessam  
APPLICANT: Ben-Arie, Nissim  
TITLE OF INVENTION: Compositions and Methods for Therapeutic Use of Atonal-association  
FILE REFERENCE: P01899US2  
CURRENT APPLICATION NUMBER: US/09/585,645A  
CURRENT FILING DATE: 2000-06-01  
PRIOR APPLICATION NUMBER: US 60/137,060  
PRIOR FILING DATE: 1999-06-01  
PRIOR APPLICATION NUMBER: US 60/176,993  
PRIOR FILING DATE: 2000-01-19  
NUMBER OF SEQ ID NOS: 70  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 54  
LENGTH: 45  
TYPE: PRT  
ORGANISM: BEETLE  
US-09-585-645A-54

Query Match 100.0%; Score 18; DB 5; Length 45;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
||||  
DB 26 SLDA 29

RESULT 4  
PCT-US02-32727-24214  
Sequence 24214, Application PC/TUS0232727  
GENERAL INFORMATION:  
APPLICANT: Mitcham, Jennifer  
APPLICANT: Skelky, Yasir  
APPLICANT: Persing, David  
APPLICANT: Bhatia, Ajay  
APPLICANT: Maisonneuve, Jean Francois  
APPLICANT: Zhang, Yanni  
APPLICANT: Wang, Sigang  
APPLICANT: Jen, Shyian  
APPLICANT: Lodes, Michael  
APPLICANT: Benson, Darin  
APPLICANT: Jones, Robert  
APPLICANT: Carter, Darriek  
APPLICANT: Barth, Brenda  
APPLICANT: Douglass, John  
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne  
FILE REFERENCE: 210121.514C1  
CURRENT APPLICATION NUMBER: PCT/US02/32727  
CURRENT FILING DATE: 2002-10-11  
NUMBER OF SEQ ID NOS: 30992  
SEQ ID NO 24214  
LENGTH: 50  
TYPE: PRT  
ORGANISM: Propionibacterium acnes  
PCT-US02-32727-24214

Query Match 100.0%; Score 18; DB 1; Length 50;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
||||  
DB 5 SLDA 8

RESULT 5  
US-10-057-498-24214

```
; Sequence 24214, Application US/10057498
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514
; CURRENT APPLICATION NUMBER: US/10/057,498
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 29212
; SEQ ID NO 24214
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Propionl acnes
; US-10-057-498-24214
```

```
Query Match
Best Local Similarity 100.0%; Score 18; DB 6; Length 50;
Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 SLDA 4
    ||||
Db 5 SLDA 8
```

```
RESULT 6
US-09-513-999C-5438
; Sequence 5438, Application US/09513999C
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.052.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 5438
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 33
; OTHER INFORMATION: Xaa-His or Asn
; US-09-513-999C-5438
```

```
Query Match
Best Local Similarity 100.0%; Score 18; DB 5; Length 52;
Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 SLDA 4
    ||||
Db 12 SLDA 15
```

```
RESULT 7
PCT-US02-32727-22975
; Sequence 22975, Application PC/TUS0232727
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siqing
; APPLICANT: Jen, Shyian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
```

```
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darrick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 22975
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Propionl acnes
; PCT-US02-32727-22975
```

```
Query Match
Best Local Similarity 100.0%; Score 18; DB 1; Length 53;
Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 SLDA 4
    ||||
Db 18 SLDA 21
```

```
RESULT 8
US-10-057-498-22975
; Sequence 22975, Application US/10057498
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514
; CURRENT APPLICATION NUMBER: US/10/057,498
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 29212
; SEQ ID NO 22975
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Propionl acnes
; US-10-057-498-22975
```

```
Query Match
Best Local Similarity 100.0%; Score 18; DB 6; Length 53;
Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 SLDA 4
    ||||
Db 18 SLDA 21
```

```
RESULT 9
PCT-US02-32727-17389
; Sequence 17389, Application PC/TUS0232727
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siqing
; APPLICANT: Jen, Shyian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darrick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
```

```

; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 17389
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Propioni acnes
PCT-US02-32727-17389
```

```
Query Match
Best Local Similarity 100.0%; Score 18; DB 1; Length 54;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 SLDA 4
    ||||
Db 28 SLDA 31
```

```

RESULT 10
; Sequence 17389, Application US/10057498
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes
; FILE REFERENCE: 210121.514
; CURRENT APPLICATION NUMBER: US/10/057,498
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 29212
; SEQ ID NO 17389
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Propioni acnes
US-10-057-498-17389
```

```
Query Match
Best Local Similarity 100.0%; Score 18; DB 6; Length 54;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 SLDA 4
    ||||
Db 28 SLDA 31
```

```

RESULT 11
; US-10-203-138A-11775
; Sequence 11775, Application US/10203138A
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: PB 0004 WO 8
; CURRENT APPLICATION NUMBER: US/10/203,138A
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 15438
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
```

```

; SEQ ID NO 11775
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009743.1
; FEATURE:
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.4
; FEATURE:
; OTHER INFORMATION: SWISSPROT HIT: P09095, EVALU 3.40e+00
; FEATURE:
; OTHER INFORMATION: EST_HUMAN HIT: R09371.1, EVALU 1.30e-02
US-10-203-138A-11775
```

```
Query Match
Best Local Similarity 100.0%; Score 18; DB 6; Length 55;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 SLDA 4
    ||||
Db 28 SLDA 31
```

```

RESULT 12
; PCT-US02-32727-2538
; Sequence 2538, Application PC/TUS0232727
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siqing
; APPLICANT: Jen, Shyian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darrick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 2538
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Propioni acnes
PCT-US02-32727-2538
```

```
Query Match
Best Local Similarity 100.0%; Score 18; DB 1; Length 56;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 SLDA 4
    ||||
Db 36 SLDA 39
```

```

RESULT 13
; PCT-US02-32727-27113
; Sequence 27113, Application PC/TUS0232727
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siqing
; APPLICANT: Jen, Shyian
```

APPLICANT: Lodes, Michael  
APPLICANT: Benson, Darin  
APPLICANT: Jones, Robert  
APPLICANT: Carter, Darrick  
APPLICANT: Barth, Brenda  
APPLICANT: Douglass, John  
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes  
FILE REFERENCE: 210121.514C1  
CURRENT APPLICATION NUMBER: PCT/US02/32727  
CURRENT FILING DATE: 2002-10-11  
NUMBER OF SEQ ID NOS: 30992  
SEQ ID NO 27113  
LENGTH: 56  
TYPE: PRT  
ORGANISM: Propionl acnes  
PCT-US02-32727-27113

Query Match 100.0%; Score 18; DB 1; Length 56;  
Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
1111  
DB 36 SLDA 39

RESULT 14  
US-10-057-498-2538  
Sequence 2538, Application US/10057498  
GENERAL INFORMATION:  
APPLICANT: Mitcham, Jennifer  
APPLICANT: Skeiky, Yasir  
APPLICANT: Persing, David  
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes  
FILE REFERENCE: 210121.514  
CURRENT APPLICATION NUMBER: US/10/057,498  
CURRENT FILING DATE: 2001-04-20  
NUMBER OF SEQ ID NOS: 29212  
SEQ ID NO 2538  
LENGTH: 56  
TYPE: PRT  
ORGANISM: Propionl acnes  
US-10-057-498-2538

Query Match 100.0%; Score 18; DB 6; Length 56;  
Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
1111  
DB 36 SLDA 39

RESULT 15  
US-10-057-498-27113  
Sequence 27113, Application US/10057498  
GENERAL INFORMATION:  
APPLICANT: Mitcham, Jennifer  
APPLICANT: Skeiky, Yasir  
APPLICANT: Persing, David  
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes  
FILE REFERENCE: 210121.514  
CURRENT APPLICATION NUMBER: US/10/057,498  
CURRENT FILING DATE: 2001-04-20  
NUMBER OF SEQ ID NOS: 29212  
SEQ ID NO 27113  
LENGTH: 56  
TYPE: PRT  
ORGANISM: Propionl acnes  
US-10-057-498-27113

Query Match 100.0%; Score 18; DB 6; Length 56;  
Best Local Similarity 100.0%; Pred. No. 3.8e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SLDA 4  
1111  
DB 36 SLDA 39

RESULT 16  
PCT-US02-32727-23045  
Sequence 23045, Application PC/TUS0232727  
GENERAL INFORMATION:  
APPLICANT: Mitcham, Jennifer  
APPLICANT: Skeiky, Yasir  
APPLICANT: Persing, David  
APPLICANT: Bhatia, Ajay  
APPLICANT: Maisonneuve, Jean Francois  
APPLICANT: Zhang, Yanni  
APPLICANT: Wang, Sigling  
APPLICANT: Jen, Shylian  
APPLICANT: Lodes, Michael  
APPLICANT: Benson, Darin  
APPLICANT: Jones, Robert  
APPLICANT: Carter, Darrick  
APPLICANT: Barth, Brenda  
APPLICANT: Douglass, John  
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne  
FILE REFERENCE: 210121.514C1  
CURRENT APPLICATION NUMBER: PCT/US02/32727  
CURRENT FILING DATE: 2002-10-11  
NUMBER OF SEQ ID NOS: 30992  
SEQ ID NO 23045  
LENGTH: 57  
TYPE: PRT  
ORGANISM: Propionl acnes  
PCT-US02-32727-23045

Query Match 100.0%; Score 18; DB 1; Length 57;  
Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
1111  
DB 22 SLDA 25

RESULT 17  
PCT-US02-32727-27575  
Sequence 27575, Application PC/TUS0232727  
GENERAL INFORMATION:  
APPLICANT: Mitcham, Jennifer  
APPLICANT: Skeiky, Yasir  
APPLICANT: Persing, David  
APPLICANT: Bhatia, Ajay  
APPLICANT: Maisonneuve, Jean Francois  
APPLICANT: Zhang, Yanni  
APPLICANT: Wang, Sigling  
APPLICANT: Jen, Shylian  
APPLICANT: Lodes, Michael  
APPLICANT: Benson, Darin  
APPLICANT: Jones, Robert  
APPLICANT: Carter, Darrick  
APPLICANT: Barth, Brenda  
APPLICANT: Douglass, John  
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne  
FILE REFERENCE: 210121.514C1  
CURRENT APPLICATION NUMBER: PCT/US02/32727  
CURRENT FILING DATE: 2002-10-11  
NUMBER OF SEQ ID NOS: 30992  
SEQ ID NO 27575  
LENGTH: 57  
TYPE: PRT  
ORGANISM: Propionl acnes  
FEATURE:

NAME/KEY: unsure  
LOCATION: (3)  
OTHER INFORMATION: Xaa = Any Amino Acid  
PCT-US02-32727-27575

Query Match 100.0%; Score 18; DB 1; Length 57;  
Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|||||  
Db 54 SLDA 57

RESULT 18  
US-10-057-498-23045  
Sequence 23045, Application US/10057498  
GENERAL INFORMATION:  
APPLICANT: Mitcham, Jennifer  
APPLICANT: Skeiky, Yasir  
APPLICANT: Persing, David  
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes  
FILE REFERENCE: 210121.514  
CURRENT APPLICATION NUMBER: US/10/057,498  
CURRENT FILING DATE: 2001-04-20  
NUMBER OF SEQ ID NOS: 29212  
SEQ ID NO 23045  
LENGTH: 57  
TYPE: PRT  
ORGANISM: Propionl acnes  
US-10-057-498-23045

Query Match 100.0%; Score 18; DB 6; Length 57;  
Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|||||  
Db 22 SLDA 25

RESULT 19  
US-10-057-498-27575  
Sequence 27575, Application US/10057498  
GENERAL INFORMATION:  
APPLICANT: Mitcham, Jennifer  
APPLICANT: Skeiky, Yasir  
APPLICANT: Persing, David  
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes  
FILE REFERENCE: 210121.514  
CURRENT APPLICATION NUMBER: US/10/057,498  
CURRENT FILING DATE: 2001-04-20  
NUMBER OF SEQ ID NOS: 29212  
SEQ ID NO 27575  
LENGTH: 57  
TYPE: PRT  
ORGANISM: Propionl acnes  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (3)  
OTHER INFORMATION: Xaa = Any Amino Acid  
US-10-057-498-27575

Query Match 100.0%; Score 18; DB 6; Length 57;  
Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|||||  
Db 54 SLDA 57

RESULT 20

PCT-US02-32727-16333  
Sequence 16333, Application PC/TUS0232727  
GENERAL INFORMATION:  
APPLICANT: Mitcham, Jennifer  
APPLICANT: Skeiky, Yasir  
APPLICANT: Persing, David  
APPLICANT: Bhatia, Ajay  
APPLICANT: Maisonneuve, Jean Francois  
APPLICANT: Zhang, Yanni  
APPLICANT: Wang, Siqing  
APPLICANT: Jen, Shyian  
APPLICANT: Iodes, Michael  
APPLICANT: Benson, Darin  
APPLICANT: Jones, Robert  
APPLICANT: Carter, Darrick  
APPLICANT: Barth, Brenda  
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne  
FILE REFERENCE: 210121.514C1  
CURRENT APPLICATION NUMBER: PCT/US02/32727  
CURRENT FILING DATE: 2002-10-11  
NUMBER OF SEQ ID NOS: 30992  
SEQ ID NO 16333  
LENGTH: 59  
TYPE: PRT  
ORGANISM: Propionl acnes  
PCT-US02-32727-16333

Query Match 100.0%; Score 18; DB 1; Length 59;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|||||  
Db 28 SLDA 31

RESULT 21  
PCT-US02-32727-28196  
Sequence 28196, Application PC/TUS0232727  
GENERAL INFORMATION:  
APPLICANT: Mitcham, Jennifer  
APPLICANT: Skeiky, Yasir  
APPLICANT: Persing, David  
APPLICANT: Bhatia, Ajay  
APPLICANT: Maisonneuve, Jean Francois  
APPLICANT: Zhang, Yanni  
APPLICANT: Wang, Siqing  
APPLICANT: Jen, Shyian  
APPLICANT: Iodes, Michael  
APPLICANT: Benson, Darin  
APPLICANT: Jones, Robert  
APPLICANT: Carter, Darrick  
APPLICANT: Barth, Brenda  
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne  
FILE REFERENCE: 210121.514C1  
CURRENT APPLICATION NUMBER: PCT/US02/32727  
CURRENT FILING DATE: 2002-10-11  
NUMBER OF SEQ ID NOS: 30992  
SEQ ID NO 28196  
LENGTH: 59  
TYPE: PRT  
ORGANISM: Propionl acnes  
PCT-US02-32727-28196

Query Match 100.0%; Score 18; DB 1; Length 59;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|||||  
Db 28 SLDA 31

RESULT 22  
PCT-US02-32727-28948  
; Sequence 28948, Application PC/TUS0232727  
; GENERAL INFORMATION:  
; APPLICANT: Mitcham, Jennifer  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Persing, David  
; APPLICANT: Bhatia, Ajay  
; APPLICANT: Maisonneuve, Jean Francois  
; APPLICANT: Zhang, Yanni  
; APPLICANT: Wang, Sigling  
; APPLICANT: Jen, Shylian  
; APPLICANT: Lodes, Michael  
; APPLICANT: Benson, Darin  
; APPLICANT: Jones, Robert  
; APPLICANT: Carter, Darrick  
; APPLICANT: Barth, Brenda  
; APPLICANT: Douglass, John  
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne  
; FILE REFERENCE: 210121.514C1  
; CURRENT APPLICATION NUMBER: PCT/US02/32727  
; CURRENT FILING DATE: 2002-10-11  
; NUMBER OF SEQ ID NOS: 30992  
; SEQ ID NO 28948  
; LENGTH: 59  
; TYPE: PRT  
; ORGANISM: P. Acnes  
PCT-US02-32727-28948

Query Match 100.0%; Score 18; DB 1; Length 59;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|||  
Db 35 SLDA 38

RESULT 23  
US-10-057-498-16333  
; Sequence 16333, Application US/10057498  
; GENERAL INFORMATION:  
; APPLICANT: Mitcham, Jennifer  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Persing, David  
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne  
; FILE REFERENCE: 210121.514  
; CURRENT APPLICATION NUMBER: US/10/057,498  
; CURRENT FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 29212  
; SEQ ID NO 16333  
; LENGTH: 59  
; TYPE: PRT  
; ORGANISM: Propionl acnes  
US-10-057-498-16333

Query Match 100.0%; Score 18; DB 6; Length 59;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|||  
Db 28 SLDA 31

RESULT 24  
US-10-057-498-28196  
; Sequence 28196, Application US/10057498  
; GENERAL INFORMATION:  
; APPLICANT: Mitcham, Jennifer  
; APPLICANT: Skeiky, Yasir

; APPLICANT: Persing, David  
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne  
; FILE REFERENCE: 210121.514  
; CURRENT APPLICATION NUMBER: US/10/057,498  
; CURRENT FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 29212  
; SEQ ID NO 28196  
; LENGTH: 59  
; TYPE: PRT  
; ORGANISM: Propionl acnes  
US-10-057-498-28196

Query Match 100.0%; Score 18; DB 6; Length 59;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|||  
Db 28 SLDA 31

RESULT 25  
US-10-057-498-28948  
; Sequence 28948, Application US/10057498  
; GENERAL INFORMATION:  
; APPLICANT: Mitcham, Jennifer  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Persing, David  
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne  
; FILE REFERENCE: 210121.514  
; CURRENT APPLICATION NUMBER: US/10/057,498  
; CURRENT FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 29212  
; SEQ ID NO 28948  
; LENGTH: 59  
; TYPE: PRT  
; ORGANISM: P. Acnes  
US-10-057-498-28948

Query Match 100.0%; Score 18; DB 6; Length 59;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|||  
Db 35 SLDA 38

RESULT 26  
PCT-US02-32727-21639  
; Sequence 21639, Application PC/TUS0232727  
; GENERAL INFORMATION:  
; APPLICANT: Mitcham, Jennifer  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Persing, David  
; APPLICANT: Bhatia, Ajay  
; APPLICANT: Maisonneuve, Jean Francois  
; APPLICANT: Zhang, Yanni  
; APPLICANT: Wang, Sigling  
; APPLICANT: Jen, Shylian  
; APPLICANT: Lodes, Michael  
; APPLICANT: Benson, Darin  
; APPLICANT: Jones, Robert  
; APPLICANT: Carter, Darrick  
; APPLICANT: Barth, Brenda  
; APPLICANT: Douglass, John  
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne  
; FILE REFERENCE: 210121.514C1  
; CURRENT APPLICATION NUMBER: PCT/US02/32727  
; CURRENT FILING DATE: 2002-10-11  
; NUMBER OF SEQ ID NOS: 30992  
; SEQ ID NO 21639  
; LENGTH: 61

TYPE: PRT  
ORGANISM: Propionl acnes  
PCT-US02-32727-21639

Query Match 100.0%; Score 18; DB 1; Length 61;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
DB 12 SLDA 15

RESULT 27

US-10-057-498-21639  
Sequence 21639, Application US/10057498

GENERAL INFORMATION:

APPLICANT: Mitcham, Jennifer

APPLICANT: Skelky, Yasir

APPLICANT: Persing, David

TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes

FILE REFERENCE: 210121.514

CURRENT APPLICATION NUMBER: US/10-057,498

CURRENT FILING DATE: 2001-04-20

NUMBER OF SEQ ID NOS: 29212

SEQ ID NO: 21639

LENGTH: 61

TYPE: PRT

ORGANISM: Propionl acnes

US-10-057-498-21639

Query Match 100.0%; Score 18; DB 6; Length 61;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
DB 12 SLDA 15

RESULT 28

US/09/513  
Sequence 7714, Application US/09513999C

GENERAL INFORMATION:

APPLICANT: Dumas Mline Edwards, J.B.

APPLICANT: Duchert, A.

APPLICANT: Giordano, J.Y.

TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

FILE REFERENCE: 59.US2.REG

CURRENT APPLICATION NUMBER: US/09/513,999C

CURRENT FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/122,487

PRIOR FILING DATE: 1999-02-26

NUMBER OF SEQ ID NOS: 36681

SOFTWARE: Patent.pm

SEQ ID NO: 7714

LENGTH: 62

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: UNSURE

LOCATION: 55

OTHER INFORMATION: Xaa= \* or Cys or Phe or His or Ile or Lys or Leu or Met or Asn C

Query Match 100.0%; Score 18; DB 5; Length 62;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
DB 59 SLDA 62

RESULT 29  
PCT-US02-32727-6416  
Sequence 6416, Application PC/TUS0232727

GENERAL INFORMATION:

APPLICANT: Mitcham, Jennifer

APPLICANT: Skelky, Yasir

APPLICANT: Persing, David

APPLICANT: Bhatia, Ajay

APPLICANT: Maisonneuve, Jean Francois

APPLICANT: Zhang, Yanni

APPLICANT: Wang, Siding

APPLICANT: Jen, Shyian

APPLICANT: Lodes, Michael

APPLICANT: Benson, Darin

APPLICANT: Jones, Robert

APPLICANT: Carter, Darrick

APPLICANT: Barth, Brenda

APPLICANT: Douglas, John

TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne

FILE REFERENCE: 210121.514c1

CURRENT APPLICATION NUMBER: PCT/US02/32727

CURRENT FILING DATE: 2002-10-11

NUMBER OF SEQ ID NOS: 30992

SEQ ID NO: 6416

LENGTH: 63

TYPE: PRT

ORGANISM: Propionl acnes

PCT-US02-32727-6416

Query Match 100.0%; Score 18; DB 1; Length 63;  
Best Local Similarity 100.0%; Pred. No. 4.3e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
DB 59 SLDA 62

RESULT 30

US-10-301-997-94  
Sequence 94, Application US/10301997

GENERAL INFORMATION:

APPLICANT: Holden, David W.

TITLE OF INVENTION: ANTI-BACTERIAL METHODS AND MATERIALS

NUMBER OF SEQUENCES: 106

CORRESPONDENCE ADDRESSES:

ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 233 South Wacker Drive/6300 Sears Tower

CITY: Chicago

STATE: Illinois

COUNTRY: United States of America

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/301,997

FILING DATE: 22-Nov-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/527,431

FILING DATE: <Unknown>

APPLICATION NUMBER: US/08/867,534

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Rin-Laures, Li-Hsien

REGISTRATION NUMBER: 33,547

REFERENCE/DOCKET NUMBER: 28341/33996

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 474-6300



TELEFAX: (312) 474-0448  
TELEX: (312) 474-6600  
INFORMATION FOR SEQ ID NO: 94:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 63 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 94:  
US-10-301-997-94

Query Match 100.0%; Score 18; DB 6; Length 63;  
Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
DB 31 SLDA 34

RESULT 31  
US-10-057-498-6416  
Sequence 6416, Application US/10057498  
GENERAL INFORMATION:  
APPLICANT: Mitcham, Jennifer  
APPLICANT: Skeiky, Yasir  
APPLICANT: Persing, David  
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes  
FILE REFERENCE: 210121.514  
CURRENT APPLICATION NUMBER: US/10/057,498  
CURRENT FILING DATE: 2001-04-20  
NUMBER OF SEQ ID NOS: 29212  
SEQ ID NO 6416  
LENGTH: 63  
TYPE: PRT  
ORGANISM: Propionl acnes  
US-10-057-498-6416

Query Match 100.0%; Score 18; DB 6; Length 63;  
Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
DB 59 SLDA 62

RESULT 32  
PCT-US02-32727-8518  
Sequence 8518, Application PC/TUS0232727  
GENERAL INFORMATION:  
APPLICANT: Mitcham, Jennifer  
APPLICANT: Skeiky, Yasir  
APPLICANT: Persing, David  
APPLICANT: Bhatia, Ajay  
APPLICANT: Maisonneuve, Jean Francois  
APPLICANT: Zhang, Yanni  
APPLICANT: Wang, Siqing  
APPLICANT: Jen, Shylan  
APPLICANT: Lodes, Michael  
APPLICANT: Benson, Darin  
APPLICANT: Jones, Robert  
APPLICANT: Carter, Derrick  
APPLICANT: Barth, Brenda  
APPLICANT: Douglass, John  
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes  
FILE REFERENCE: 210121.514C1  
CURRENT APPLICATION NUMBER: PCT/US02/32727  
CURRENT FILING DATE: 2002-10-11  
NUMBER OF SEQ ID NOS: 30992  
SEQ ID NO 8518  
LENGTH: 65  
TYPE: PRT

ORGANISM: Propionl acnes  
PCT-US02-32727-8518

Query Match 100.0%; Score 18; DB 1; Length 65;  
Best Local Similarity 100.0%; Pred. No. 4.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
DB 46 SLDA 49

RESULT 33  
US-10-057-498-8518  
Sequence 8518, Application US/10057498  
GENERAL INFORMATION:  
APPLICANT: Mitcham, Jennifer  
APPLICANT: Skeiky, Yasir  
APPLICANT: Persing, David  
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne  
FILE REFERENCE: 210121.514  
CURRENT APPLICATION NUMBER: US/10/057,498  
CURRENT FILING DATE: 2001-04-20  
NUMBER OF SEQ ID NOS: 29212  
SEQ ID NO 8518  
LENGTH: 65  
TYPE: PRT  
ORGANISM: Propionl acnes  
US-10-057-498-8518

Query Match 100.0%; Score 18; DB 6; Length 65;  
Best Local Similarity 100.0%; Pred. No. 4.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
DB 46 SLDA 49

RESULT 34  
PCT-US02-32727-2064  
Sequence 2064, Application PC/TUS0232727  
GENERAL INFORMATION:  
APPLICANT: Mitcham, Jennifer  
APPLICANT: Skeiky, Yasir  
APPLICANT: Persing, David  
APPLICANT: Bhatia, Ajay  
APPLICANT: Maisonneuve, Jean Francois  
APPLICANT: Zhang, Yanni  
APPLICANT: Wang, Siqing  
APPLICANT: Jen, Shylan  
APPLICANT: Lodes, Michael  
APPLICANT: Benson, Darin  
APPLICANT: Jones, Robert  
APPLICANT: Carter, Derrick  
APPLICANT: Barth, Brenda  
APPLICANT: Douglass, John  
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne  
FILE REFERENCE: 210121.514C1  
CURRENT APPLICATION NUMBER: PCT/US02/32727  
CURRENT FILING DATE: 2002-10-11  
NUMBER OF SEQ ID NOS: 30992  
SEQ ID NO 2064  
LENGTH: 68  
TYPE: PRT  
ORGANISM: Propionl acnes  
PCT-US02-32727-2064

Query Match 100.0%; Score 18; DB 1; Length 68;  
Best Local Similarity 100.0%; Pred. No. 4.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4

Db 1 SLDA 4

RESULT 35

US-10-057-498-2064  
; Sequence 2064, Application US/10057498  
; GENERAL INFORMATION:  
; APPLICANT: Mitcham, Jennifer  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Persing, David  
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes  
; FILE REFERENCE: 210121.514  
; CURRENT APPLICATION NUMBER: US/10/057,498  
; CURRENT FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 29212  
; SEQ ID NO 2064  
; LENGTH: 68  
; TYPE: PRT  
; ORGANISM: Propionl acnes  
US-10-057-498-2064

Query Match 100.0%; Score 18; DB 6; Length 68;

Best Local Similarity 100.0%; Pred. No. 4.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4

Db 1 SLDA 4

RESULT 36

PCT-US02-32727-6955  
; Sequence 6955, Application PC/TUS0232727  
; GENERAL INFORMATION:  
; APPLICANT: Mitcham, Jennifer  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Persing, David  
; APPLICANT: Bhatia, Ajay  
; APPLICANT: Maisonneuve, Jean Francois  
; APPLICANT: Zhang, Yanni  
; APPLICANT: Wang, Siqing  
; APPLICANT: Jen, Shyian  
; APPLICANT: Lodes, Michael  
; APPLICANT: Benson, Darin  
; APPLICANT: Jones, Robert  
; APPLICANT: Carter, Darrick  
; APPLICANT: Barth, Brenda  
; APPLICANT: Douglass, John  
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes  
; FILE REFERENCE: 210121.514C1  
; CURRENT APPLICATION NUMBER: PCT/US02/32727  
; CURRENT FILING DATE: 2002-10-11  
; NUMBER OF SEQ ID NOS: 30992  
; SEQ ID NO 6955  
; LENGTH: 70  
; TYPE: PRT  
; ORGANISM: Propionl acnes  
PCT-US02-32727-6955

Query Match 100.0%; Score 18; DB 1; Length 70;

Best Local Similarity 100.0%; Pred. No. 4.8e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4

Db 35 SLDA 38

RESULT 37

PCT-US02-32727-28376  
; Sequence 28376, Application PC/TUS0232727  
; GENERAL INFORMATION:

APPLICANT: Mitcham, Jennifer

APPLICANT: Skeiky, Yasir

APPLICANT: Persing, David

APPLICANT: Bhatia, Ajay

APPLICANT: Maisonneuve, Jean Francois

APPLICANT: Zhang, Yanni

APPLICANT: Wang, Siqing

APPLICANT: Jen, Shyian

APPLICANT: Lodes, Michael

APPLICANT: Benson, Darin

APPLICANT: Jones, Robert

APPLICANT: Carter, Darrick

APPLICANT: Barth, Brenda

APPLICANT: Douglass, John

APPLICANT: Persing, David

APPLICANT: Wang, Siqing

APPLICANT: Jen, Shyian

APPLICANT: Lodes, Michael

APPLICANT: Benson, Darin

APPLICANT: Jones, Robert

APPLICANT: Carter, Darrick

APPLICANT: Barth, Brenda

APPLICANT: Douglass, John

APPLICANT: Persing, David

APPLICANT: Wang, Siqing

APPLICANT: Jen, Shyian

APPLICANT: Lodes, Michael

APPLICANT: Benson, Darin

APPLICANT: Jones, Robert

APPLICANT: Carter, Darrick

APPLICANT: Barth, Brenda

APPLICANT: Douglass, John

APPLICANT: Persing, David

APPLICANT: Wang, Siqing

APPLICANT: Jen, Shyian

APPLICANT: Lodes, Michael

APPLICANT: Benson, Darin

APPLICANT: Jones, Robert

APPLICANT: Carter, Darrick

APPLICANT: Barth, Brenda

APPLICANT: Douglass, John

APPLICANT: Persing, David

APPLICANT: Wang, Siqing

APPLICANT: Jen, Shyian

APPLICANT: Lodes, Michael

APPLICANT: Benson, Darin

APPLICANT: Jones, Robert

APPLICANT: Carter, Darrick

APPLICANT: Barth, Brenda

APPLICANT: Douglass, John

APPLICANT: Persing, David

APPLICANT: Wang, Siqing

RESULT 38

US-10-057-498-6955

Sequence 6955, Application US/10057498

GENERAL INFORMATION:

APPLICANT: Mitcham, Jennifer

APPLICANT: Skeiky, Yasir

APPLICANT: Persing, David

TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne

FILE REFERENCE: 210121.514

RESULT 39

US-10-057-498-28376

Sequence 28376, Application US/10057498

GENERAL INFORMATION:

APPLICANT: Mitcham, Jennifer

APPLICANT: Skeiky, Yasir

APPLICANT: Persing, David

TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne

FILE REFERENCE: 210121.514

FILE REFERENCE: 210121.514  
CURRENT APPLICATION NUMBER: US/10/057,498  
CURRENT FILING DATE: 2001-04-20  
NUMBER OF SEQ ID NOS: 29212  
SEQ ID NO 28376  
LENGTH: 70  
TYPE: PRT  
ORGANISM: Propionibacterium  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (40)  
OTHER INFORMATION: Xaa - Any Amino Acid  
US-10-057-498-28376

Query Match 100.0%; Score 18; DB 6; Length 70;  
Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
1111  
DB 24 SLDA 27

RESULT 40  
US-10-203-138A-11100  
GENERAL INFORMATION:  
APPLICANT: Molecular Dynamics, Inc.  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: PB 0004 WO 8  
CURRENT FILING DATE: 2002-08-02  
PRIOR APPLICATION NUMBER: US/10/203,138A  
PRIOR FILING DATE: 04 February 2000 (04.02.00)  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 26 May 2000 (26.05.00)  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 03 August 2000 (03.08.00)  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 03 October 2000 (03.10.00)  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 27 September 2000 (27.09.00)  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 21 September 2000 (21.09.00)  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 30 June 2000 (30.06.00)  
NUMBER OF SEQ ID NOS: 13438  
SOFTWARE: Molecular Dynamics Sequence Listing Engine  
SEQ ID NO 11100  
LENGTH: 70  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO 283845.14  
FEATURE:  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3  
FEATURE:  
OTHER INFORMATION: EST\_HUMAN HIT: BE746542.1, EVALUATE 5.00e-35  
FEATURE:  
OTHER INFORMATION: SWISSPROT HIT: Q15750, EVALUATE 4.00e-36  
US-10-203-138A-11100

Query Match 100.0%; Score 18; DB 6; Length 70;  
Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
1111

DB 11 SLDA 14

RESULT 41  
PCT-US02-32727-13114  
Sequence 13114, Application PC/TUS0232727  
GENERAL INFORMATION:  
APPLICANT: Mitcham, Jennifer  
APPLICANT: Skelky, Yasir  
APPLICANT: Persing, David  
APPLICANT: Bhatia, Ajay  
APPLICANT: Maisonneuve, Jean Francois  
APPLICANT: Zhang, Yanni  
APPLICANT: Wang, Siqing  
APPLICANT: Jen, Shyian  
APPLICANT: Lodes, Michael  
APPLICANT: Benson, Darin  
APPLICANT: Jones, Robert  
APPLICANT: Carter, Darrick  
APPLICANT: Barth, Brenda  
APPLICANT: Douglass, John  
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne  
FILE REFERENCE: 210121.514C1  
CURRENT APPLICATION NUMBER: PCT/US02/32727  
CURRENT FILING DATE: 2002-10-11  
NUMBER OF SEQ ID NOS: 30992  
SEQ ID NO 13114  
LENGTH: 72  
TYPE: PRT  
ORGANISM: Propionibacterium  
PCT-US02-32727-13114

Query Match 100.0%; Score 18; DB 1; Length 72;  
Best Local Similarity 100.0%; Pred. No. 5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
1111  
DB 66 SLDA 69

RESULT 42  
PCT-US02-32727-27093  
Sequence 27093, Application PC/TUS0232727  
GENERAL INFORMATION:  
APPLICANT: Mitcham, Jennifer  
APPLICANT: Skelky, Yasir  
APPLICANT: Persing, David  
APPLICANT: Bhatia, Ajay  
APPLICANT: Maisonneuve, Jean Francois  
APPLICANT: Zhang, Yanni  
APPLICANT: Wang, Siqing  
APPLICANT: Jen, Shyian  
APPLICANT: Lodes, Michael  
APPLICANT: Benson, Darin  
APPLICANT: Jones, Robert  
APPLICANT: Carter, Darrick  
APPLICANT: Barth, Brenda  
APPLICANT: Douglass, John  
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne  
FILE REFERENCE: 210121.514C1  
CURRENT APPLICATION NUMBER: PCT/US02/32727  
CURRENT FILING DATE: 2002-10-11  
NUMBER OF SEQ ID NOS: 30992  
SEQ ID NO 27093  
LENGTH: 72  
TYPE: PRT  
ORGANISM: Propionibacterium  
PCT-US02-32727-27093

Query Match 100.0%; Score 18; DB 1; Length 72;  
Best Local Similarity 100.0%; Pred. No. 5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
Db 66 SLDA 69

RESULT 43  
US-10-264-213-210

; Sequence 210, Application US/10264213  
; GENERAL INFORMATION:  
; APPLICANT: Glenn, Matthew  
; APPLICANT: Havukkala, Ilkka J  
; APPLICANT: Lubbers, Mark William  
; APPLICANT: Dekker, James  
; TITLE OF INVENTION: Polynucleotides, materials incorporating  
; FILE REFERENCE: 11000.1043c3  
; CURRENT APPLICATION NUMBER: US/10/264,213  
; CURRENT FILING DATE: 2002-10-03  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO: 210  
; LENGTH: 72  
; TYPE: PRT  
; ORGANISM: Lactobacillus rhamnosus  
US-10-264-213-210

Query Match 100.0%; Score 18; DB 6; Length 72;  
Best Local Similarity 100.0%; Pred. No. 5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
Db 14 SLDA 17

RESULT 44

US-10-057-498-13114  
; Sequence 13114, Application US/10057498  
; GENERAL INFORMATION:  
; APPLICANT: Mitcham, Jennifer  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Persing, David  
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes  
; FILE REFERENCE: 210121.514  
; CURRENT APPLICATION NUMBER: US/10/057,498  
; CURRENT FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 29212  
; SEQ ID NO: 13114  
; LENGTH: 72  
; TYPE: PRT  
; ORGANISM: Propionibacterium acnes  
US-10-057-498-13114

Query Match 100.0%; Score 18; DB 6; Length 72;  
Best Local Similarity 100.0%; Pred. No. 5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
Db 66 SLDA 69

RESULT 45

US-10-057-498-27093  
; Sequence 27093, Application US/10057498  
; GENERAL INFORMATION:  
; APPLICANT: Mitcham, Jennifer  
; APPLICANT: Skeiky, Yasir  
; APPLICANT: Persing, David  
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes  
; FILE REFERENCE: 210121.514  
; CURRENT APPLICATION NUMBER: US/10/057,498

; CURRENT FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 29212  
; SEQ ID NO: 27093  
; LENGTH: 72  
; TYPE: PRT  
; ORGANISM: Propionibacterium acnes  
US-10-057-498-27093

Query Match 100.0%; Score 18; DB 6; Length 72;  
Best Local Similarity 100.0%; Pred. No. 5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
Db 66 SLDA 69

Search completed: February 6, 2003, 11:39:03  
Job time : 15 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: February 6, 2003, 11:16:04 ; Search time 10.5 Seconds  
(without alignments)  
36.623 Million cell updates/sec

Title: PAT943-4  
Perfect score: 18  
Sequence: 1 slide 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 2413

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73:\*

1: PIR1:.\*  
2: PIR2:.\*  
3: PIR3:.\*  
4: PIR4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	100.0	20	2 A44921	hydroxypruvate re
2	18	100.0	33	2 G86126	hypothetical prote
3	18	100.0	42	2 D49410	t-complex polypept
4	18	100.0	44	2 E85608	hypothetical prote
5	18	100.0	45	1 HPBO	haptoglobin precur
6	18	100.0	49	2 H70537	hypothetical prote
7	18	100.0	52	2 B49529	RNA-directed RNA p
8	18	100.0	53	2 AC2486	hypothetical prote
9	18	100.0	55	2 AB0174	probable ribosome
10	18	100.0	56	2 E86594	hypothetical prote
11	18	100.0	56	2 A72031	hypothetical prote
12	18	100.0	62	2 F81541	hypothetical prote
13	18	100.0	63	2 A43397	4-oxalocrotonate t
14	18	100.0	63	2 S24422	4-oxalocrotonate t
15	18	100.0	64	2 F84858	haptoglobin precur
16	18	100.0	73	1 HPRB	hypothetical prote
17	18	100.0	73	2 E70523	exodeoxyribonuclea
18	18	100.0	75	2 A81539	thioredoxin relate
19	18	100.0	76	2 G83935	hypothetical prote
20	18	100.0	85	2 A69435	gene 55.1 protein
21	18	100.0	87	2 ZGBPT9	hypothetical prote
22	18	100.0	87	2 E84162	hypothetical prote
23	18	100.0	88	2 B97373	conserved hypotnet
24	18	100.0	88	2 AB2591	DNA-binding protei
25	18	100.0	90	2 T44780	embryonic abundant
26	18	100.0	91	2 T09285	hypothetical prote
27	18	100.0	91	2 T09293	conserved hypotnet
28	18	100.0	92	2 E81209	ribosomal protein
29	18	100.0	95	2 T10931	

30	18	100.0	96	2 E81786	conserved hypotnet
31	18	100.0	96	2 T12686	hypothetical prote
32	18	100.0	98	2 JC2403	PKS8 homolog misma
33	18	100.0	99	2 S70645	cell division cont
34	18	100.0	99	2 AC1127	hypothetical prote
35	18	100.0	99	2 AH1487	hypothetical prote
36	18	100.0	101	2 T35232	hypothetical prote
37	18	100.0	102	2 S19225	embryonic abundant
38	18	100.0	104	2 T41937	hypothetical prote
39	18	100.0	107	2 H87712	phosphoribosyl-ATP
40	18	100.0	107	2 T17856	hypothetical prote
41	18	100.0	108	2 T06898	hypothetical prote
42	18	100.0	108	2 T34677	hypothetical prote
43	18	100.0	109	2 H72641	hypothetical prote
44	18	100.0	110	2 D72019	sigma regulatory f
45	18	100.0	110	2 C86604	sigma regulatory f

## ALIGNMENTS

RESULT 1  
A44921  
hydroxypruvate reductase (EC 1.1.1.81) - Methylobacterium extorquens (fragment)  
C:Species: Methylobacterium extorquens  
C>Date: 01-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 08-Oct-1999  
C:Accession: A44921  
R:Chistoserdova, L.V.; Lidstrom, M.E.  
J. Bacteriol. 174, 71-77, 1992  
A>Title: Cloning, mutagenesis, and physiological effect of a hydroxypruvate reductas  
A:Reference number: A44921; MUID:92104992; PMID:1729225  
A:Contents: AM1  
A:Accession: A44921  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-20 <CH1>  
A:Cross-references: GB:M81443; NID:9150010; PIDN:AA25378.1; PID:9150011  
A>Note: sequence extracted from NCBI backbone (NCBI:75202, NCBI:P75203)  
C:Keywords: oxidoreductase

Query Match 100.0%; Score 18; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 63;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
DB 12 SLDA 15

RESULT 2  
G86126  
hypothetical protein Z5889 [Imported] - Escherichia coli (strain O157:H7, substrain E  
C:Species: Escherichia coli  
C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C:Accession: G86126  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May  
hiller, L.; Grothbeck, E.J.; Davis, N.W.; Lim, A.; Diallanata, E.; Potamousis, K.; Apoda  
Nature 409, 529-533, 2001  
A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: G86126  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-33 <STO>  
A:Cross-references: GB:AE005174; NID:912519291; PIDN:AA659475.1; GSPDB:GN00145; UWGP:  
A:Experimental source: strain O157:H7, substrain ED933  
C:Genetics:  
A:Gene: Z5889

Query Match 100.0%; Score 18; DB 2; Length 33;  
Best Local Similarity 100.0%; Pred. No. 11e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|||||  
Db 6 SLDA 9

## RESULT 3

D49410  
t-complex polypeptide 1 homolog (peak 1 fraction) - rabbit (fragments)  
N:Alternate names: chaparonin homolog (peak 1)  
C:Species: *Oryctolagus cuniculus* (domestic rabbit)  
C:Date: 21-Sep-1994 #sequence\_revision 18-Nov-1994 #text\_change 23-Mar-1995  
C:Accession: D49410  
R:Rommelaere, H.; Van Troy, M.; Gao, Y.; Melki, R.; Cowan, N.J.; Vandekerckhove, J.; Am  
Proc. Natl. Acad. Sci. U.S.A. 90, 11975-11979, 1993  
A:Title: Eukaryotic cytosolic chaparonin contains t-complex polypeptide 1 and seven rela  
A:Reference number: A49410; MUID:94089752; PMID:7903435  
A:Accession: D49410  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-42 <ROM>  
A:Experimental source: reticulocyte  
A:Note: sequence modified after extraction from NCBI backbone (NCBIP:141042)  
A:Note: sequence extracted from NCBI backbone (NCBIP:141042)

Query Match 100.0%; Score 18; DB 2; Length 42;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|||||  
Db 19 SLDA 22

## RESULT 4

E85608  
hypothetical protein 21184 [imported] - *Escherichia coli* (strain 0157:H7, substrain EDL9  
C:Species: *Escherichia coli*  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C:Accession: E85608; F85659  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoudis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* 0157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: E85608  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-44 <STO>  
A:Cross-references: GB:AE005174; NID:q12513992; PIDN:AA655329.1; GSPDB:GN00145; UWGP:Z11  
A:Experimental source: strain 0157:H7, substrain EDL933  
A:Accession: F85659  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-44 <ST2>  
A:Cross-references: GB:AE005174; NID:q12514512; PIDN:AA655738.1; GSPDB:GN00145; UWGP:Z16  
A:Experimental source: strain 0157:H7, substrain EDL933  
A:Gene: Z1184; Z1623

Query Match 100.0%; Score 18; DB 2; Length 44;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|||||  
Db 10 SLDA 13

## RESULT 5

HPBO  
haptoglobin precursor - bovine (fragments)  
C:Species: *Bos primigenius taurus* (cattle)  
C:Date: 28-Feb-1992 #sequence\_revision 31-May-1996 #text\_change 31-Oct-1997

C:Accession: A40430; B40430  
R:Moriwatsu, M.; Syuto, B.; Shimada, N.; Fujinaga, T.; Yamamoto, S.; Saito, M.; Naiki  
J. Biol. Chem. 266, 11833-11837, 1991  
A:Title: Isolation and characterization of bovine haptoglobin from acute phase sera.  
A:Reference number: A40430; MUID:91268055; PMID:1904872  
A:Accession: A40430

A:Molecule type: protein  
A:Residues: 1-25 <MOR>  
A:Accession: B40430  
A:Molecule type: protein  
A:Residues: 26-45 <MO2>  
C:Comment: The mature haptoglobin molecule is a dimer of heterodimers, the two chains  
C:Comment: Haptoglobin is a plasma glycoprotein; haptoglobin forms a complex with hem  
C:Superfamily: haptoglobin; complement factor H repeat homology; trypsin homology  
C:Keywords: acute phase; glycoprotein; hemoglobin binding; heterotetramer; iron trans  
F:1-25/Product: haptoglobin alpha chain (fragment) #status experimental <ALP>  
F:26-45/Product: haptoglobin beta chain (fragment) #status experimental <BET>

Query Match 100.0%; Score 18; DB 1; Length 45;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|||||  
Db 30 SLDA 33

## RESULT 6

H70537  
hypothetical protein RV119c - *Mycobacterium tuberculosis* (strain H37RV)  
C:Species: *Mycobacterium tuberculosis*  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
C:Accession: H70537  
R:Coile, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon  
; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd,  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete geno  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: H70537  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-49 <COL>  
A:Cross-references: GB:Z95585; GB:AL123456; NID:q3261787; PIDN:CAB09051.1; PID:es1714  
A:Experimental source: strain H37RV  
A:Gene: RV119c

Query Match 100.0%; Score 18; DB 2; Length 49;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|||||  
Db 39 SLDA 42

## RESULT 7

B49529  
RNA-directed RNA polymerase (EC 2.7.7.48) - human astrovirus type 1 (fragment)  
C:Species: human astrovirus type 1  
C:Date: 07-Apr-1994 #sequence\_revision 18-Nov-1994 #text\_change 08-Dec-1994  
C:Accession: B49529  
R:Lewis, T.L.; Greenberg, H.B.; Herrmann, J.E.; Smith, L.S.; Matsui, S.M.  
J. Virol. 68, 77-83, 1994  
A:Title: Analysis of astrovirus serotype 1 RNA, identification of the viral RNA-depen  
A:Reference number: A49529; MUID:94076464; PMID:8254779  
A:Accession: B49529  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: nucleic acid  
A:Residues: 1-52 <LEW>  
A:Note: sequence extracted from NCBI backbone (NCBIP:141412)

## C:Keywords: nucleotidyltransferase

Query Match 100.0%; Score 18; DB 2; Length 52;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
 ||||  
 Db 23 SLDA 26

## RESULT 8

AC2486  
 hypothetical protein asr7067 [imported] - Nostoc sp. (strain PCC 7120) plasmid pcc7120a1

C:Species: Nostoc sp.

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C>Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002

C:Accession: AC2486

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi, Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AC2486

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-53 <KUR>

A:Cross-references: GB:BA000020; PIDN:BAF78151.1; PID:g17135605; GSPDB:GN00180

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: asr7067

A:Genome: plasmid

Query Match 100.0%; Score 18; DB 2; Length 53;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
 ||||  
 Db 43 SLDA 46

## RESULT 9

AB0174

probable ribosome modulation factor rnf [imported] - Yersinia pestis (strain CO92)

C:Species: Yersinia pestis

C>Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 02-Nov-2001

C:Accession: AB0174

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B. deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett, Nature 413, 533-537, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AB0174

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-55 <KUR>

A:Cross-references: GB:AL590842; PIDN:CAC90253.1; PID:g15979473; GSPDB:GN00175

C:Genetics:

A:Gene: rnf

Query Match 100.0%; Score 18; DB 2; Length 55;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
 ||||  
 Db 33 SLDA 36

RESULT 10  
 E86594

hypothetical protein CPJ0831 [imported] - Chlamydia pneumoniae (strain J138)

C:Species: Chlamydia pneumoniae

C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 02-Mar-2001

C:Accession: E86594

R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.;

Nucleic Acids Res. 28, 2311-2314, 2000

A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.

A:Reference number: AB6491; MUID:20330349; PMID:10871362

A:Accession: E86594

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-56 <STO>

A:Cross-references: GB:BA000008; NID:g8979205; PIDN:BAA99039.1; GSPDB:GN00142

A:Experimental source: strain J138

C:Genetics:

A:Gene: CPJ0831

Query Match 100.0%; Score 18; DB 2; Length 56;  
 Best Local Similarity 100.0%; Pred. No. 2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
 ||||  
 Db 47 SLDA 50

## RESULT 11

AT2031

hypothetical protein CP1039 [imported] - Chlamydia pneumoniae (strains CWL029 and

C:Species: Chlamydia pneumoniae

C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 11-May-2000

C:Accession: A72031; D81509

R:Katzman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, Nature Genet. 21, 385-389, 1999

A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.

A:Reference number: A72000; MUID:99206606; PMID:10192388

A:Accession: A72031

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-56 <ARN>

A:Cross-references: GB:AE001663; GB:AE001363; NID:g4377134; PIDN:AAD18968.1; PID:g437

A:Experimental source: strain CWL029

R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwin, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzbe

Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39

A:Reference number: AB1500; MUID:20150255; PMID:10684935

A:Accession: D81509

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-56 <REA>

A:Cross-references: GB:AE002261; GB:AE002161; NID:g7189950; PIDN:AAF38814.1; PID:g718

A:Experimental source: strain AR39, HL cells

C:Genetics:

A:Gene: CPN0831; CP1039

Query Match 100.0%; Score 18; DB 2; Length 56;  
 Best Local Similarity 100.0%; Pred. No. 2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
 ||||  
 Db 47 SLDA 50

## RESULT 12

F81541

hypothetical protein CP0758 [imported] - Chlamydia pneumoniae (strain AR39)

C:Species: Chlamydia pneumoniae

C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 11-May-2000

C:Accession: F81541

R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwin, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzbe

Nucleic Acids Res. 28, 1397-1406, 2000  
A:Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39.  
A:Reference number: AB1500; MUID:20150255; PMID:10684935  
A:Accession: F81541  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-62 <REA>  
A:Cross-references: GB:AE002235; GB:AE002161; NID:g7189672; PIDN:AAF38560.1; PID:g718967  
A:Experimental source: strain AR39, HL cells  
A:Genetics:

Query Match  
Best Local Similarity 100.0%; Score 18; DB 2; Length 62;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
DB 32 SLDA 35

## RESULT 13

4-oxalocrotonate tautomerase (EC 5.3.2.-) xylH - Pseudomonas putida plasmid TOL pmw0  
C:Species: Pseudomonas putida  
C:Date: 10-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 17-Mar-2000  
C:Accession: A43397; S35225  
R:Chen, L.H.; Kenyon, G.L.; Curtin, F.; Harayama, S.; Bemberek, M.E.; Hajjipour, G.; Whit  
J. Biol. Chem. 267, 17716-17721, 1992  
A:Title: 4-Oxalocrotonate tautomerase, an enzyme composed of 62 amino acid residues per  
A:Reference number: A43397; MUID:92388122; PMID:1339435  
A:Accession: A43397  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-63 <CHE>  
A:Note: sequence extracted from NCHI backbone (NCBIN:112741, NCBIP:112742)  
R:Harayama, S.; Reik, M.  
Mol. Gen. Genet. 239, 81-89, 1993  
A:Title: Comparison of the nucleotide sequences of the meta-cleavage pathway genes of TC  
cleotide substitutions contribute to enzyme evolution.  
A:Reference number: S35222; MUID:93288011; PMID:8510667  
A:Accession: S35225  
A:Molecule type: DNA  
A:Residues: 1-63 <HAR>  
A:Cross-references: EMBL:M94186; NID:g150970; PIDN:AAA25694.1; PID:g150974  
C:Genetics:  
A:Gene: xylH  
A:Genome: plasmid  
C:Superfamily: 4-oxalocrotonate tautomerase  
C:Keywords: aromatic hydrocarbon catabolism; intramolecular oxidoreductase; isomerase  
F:2-63/Product: 4-oxalocrotonate tautomerase #status predicted <MAT>  
F:2/Active site: Pro #status predicted

Query Match  
Best Local Similarity 100.0%; Score 18; DB 2; Length 63;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
DB 31 SLDA 34

## RESULT 14

4-oxalocrotonate tautomerase (EC 5.3.2.-) dmpI [similarity] - Pseudomonas putida  
C:Species: Pseudomonas putida  
C:Date: 19-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 17-Mar-2000  
C:Accession: S24422  
R:Shingler, V.; Powlowski, J.; Marklund, U.  
J. Bacteriol. 174, 711-724, 1992  
A:Title: Nucleotide sequence and functional analysis of the complete phenol/3,4-dimethyl  
A:Reference number: S24417; MUID:92121108; PMID:1733207  
A:Accession: S24422

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-63 <SHI>  
A:Cross-references: EMBL:X60835; NID:g45681; PIDN:CAA43229.1; PID:g45686  
C:Superfamily: 4-oxalocrotonate tautomerase  
C:Keywords: aromatic hydrocarbon catabolism; intramolecular oxidoreductase; isomerase  
F:2-63/Product: 4-oxalocrotonate tautomerase #status predicted <MAT>  
F:2/Active site: Pro #status predicted

Query Match  
Best Local Similarity 100.0%; Score 18; DB 2; Length 63;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
DB 31 SLDA 34

## RESULT 15

hypothetical protein At2g42820 [imported] - Arabidopsis thaliana  
F84858  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: F84858  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: F84858  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-64 <STO>  
A:Cross-references: GB:AE002093; NID:g4512672; PIDN:AAD21726.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g42820  
A:Map position: 2

Query Match  
Best Local Similarity 100.0%; Score 18; DB 2; Length 64;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
DB 33 SLDA 36

## RESULT 16

haptoglobin precursor - rabbit (fragments)  
N:Contents: haptoglobin alpha chain; haptoglobin beta chain  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 31-Mar-1988 #sequence\_revision 31-May-1996 #text\_change 11-May-2000  
C:Accession: A19376; A26503  
R:Chow, V.; Murray, R.K.; Dixon, J.D.; Kurosky, A.  
FEBS Lett. 153, 275-279, 1983  
A:Title: Biosynthesis of rabbit haptoglobin: chemical evidence for a single chain pre  
A:Reference number: A19376; MUID:84005090; PMID:6413248  
A:Accession: A19376  
A:Molecule type: protein  
A:Residues: 1-33 <CH2>  
R:Kurosky, A.; Kim, H.H.; Touchstone, B.  
Comp. Biochem. Physiol. B 55, 453-459, 1976  
A:Title: Comparative sequence analysis of the N-terminal region of rat, rabbit, and d  
A:Reference number: A90931; MUID:77025019; PMID:975782  
A:Accession: A26503  
A:Molecule type: protein  
A:Residues: 34-73 <KRP>  
C:Comment: The mature haptoglobin molecule is a dimer of heterodimers, the two chains  
C:Superfamily: haptoglobin; complement factor H repeat homology; trypsin homology  
C:Keywords: acute phase; glycoprotein; hemoglobin binding; heterotetramer; iron trans



F:1-18/Domain: signal sequence #status experimental <SIG>  
F:19-33/Product: haptoglobin alpha chain (fragment) #status experimental <ACH>  
F:34-73/Product: haptoglobin beta chain (fragment) #status experimental <BCH>  
F:34-73/Domain: trypsin homology (fragment) <TRY>

Query Match 100.0%; Score 18; DB 1; Length 73;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|||||  
DB 38 SLDA 41

RESULT 17  
E70523  
hypothetical protein RV0300 - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
C:Accession: E70523  
R:Coile, S.T.; Brosch, R.; Parhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajadream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:9825987; PMID:9634230  
A:Accession: E70523  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-73 <COL>  
A:Cross-references: GB:296800; GB:AL12456; NID:g3261800; PIDN:CAB09591.1; PID:e321658;  
A:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: RV0300

Query Match 100.0%; Score 18; DB 2; Length 73;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|||||  
DB 16 SLDA 19

RESULT 18  
AB1539  
exodeoxyribonuclease, small chain CP0788 [imported] - Chlamydomonas reinhardtii (strain A)  
C:Species: Chlamydomonas reinhardtii  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 11-May-2000  
C:Accession: AB1539  
R:Read, T.D.; Brumham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwin, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000  
A>Title: Genome sequences of Chlamydomonas reinhardtii and Chlamydomonas reinhardtii  
A:Reference number: AB1500; MUID:20150255; PMID:10684935  
A:Accession: AB1539  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-75 <REA>  
A:Cross-references: GB:AE002218; GB:AE002161; NID:g7189693; PIDN:AAF38587.1; PID:g718970  
A:Experimental source: strain AR39, HL cells  
C:Genetics:  
A:Gene: CP0788

Query Match 100.0%; Score 18; DB 2; Length 75;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|||||  
DB 26 SLDA 29

RESULT 19  
G83935  
thioredoxin related protein BH2287 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C:Accession: G83935  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; I Nucleic Acids Res. 28, 4317-4331, 2000  
A>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans  
A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: G83935  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-76 <SNO>  
A:Cross-references: GB:AP001515; GB:BA000004; NID:g10174886; PIDN:BAB06006.1; GSPDB:C  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BH2287

Query Match 100.0%; Score 18; DB 2; Length 76;  
Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|||||  
DB 55 SLDA 58

RESULT 20  
A69435  
hypothetical protein AF1482 - Archaeoglobus fulgidus  
C:Species: Archaeoglobus fulgidus  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
C:Accession: A69435  
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; De Giodok, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L. Nature 390, 364-370, 1997  
A:Authors: Overbeek, R.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, Smith, H.O.; Woese, C.R.; Venter, J.C.  
A>Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arc  
A:Reference number: A69250; MUID:98049343; PMID:9389475  
A:Accession: A69435  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-85 <KLE>  
A:Cross-references: GB:AE001000; GB:AE000782; NID:g2689323; PIDN:AAB89770.1; PID:g264

Query Match 100.0%; Score 18; DB 2; Length 85;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|||||  
DB 24 SLDA 27

RESULT 21  
ZGBP79  
gene 55.1 protein - phage T4  
C:Species: phage T4  
C:Date: 31-Dec-1989 #sequence\_revision 31-Dec-1989 #text\_change 23-Jul-1999  
C:Accession: F30292  
R:Tomaschewski, J.; Rueger, W. Nucleic Acids Res. 15, 3632-3633, 1987  
A>Title: Nucleotide sequence and primary structures of gene products coded for by the  
A:Reference number: A30291; MUID:87203398; PMID:3575111  
A:Accession: F30292  
A:Molecule type: DNA  
A:Residues: 1-87 <TOM>  
A:Cross-references: GB:Y00122; NID:g15346; PIDN:CAA68320.1; PID:g15362

C:Genetics:  
A:Gene: 55.1  
C:Superfamily: phage T4 gene 55.1 protein

Query Match  
Best Local Similarity 100.0%; Score 18; DB 1; Length 87;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SUDA 4  
||||  
DB 58 SUDA 61

## RESULT 22

E84162  
Hypothetical protein Vng0025h [imported] - Halobacterium sp. NRC-1  
C:Species: Halobacterium sp. NRC-1  
C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: E84162  
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leitauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.F.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li  
A:Title: Genome sequence of Halobacterium species NRC-1.  
A:Reference number: A84160; MID:20504483; PMID:11016950  
A:Accession: E84162  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-87 <STO>  
A:Cross-references: GB:AE004437; MID:g10579673; PIDN:AA618665.1; GSPDB:GN00138  
C:Genetics:  
A:Gene: VNG0025H

Query Match  
Best Local Similarity 100.0%; Score 18; DB 2; Length 87;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SUDA 4  
||||  
DB 9 SUDA 12

## RESULT 23

B97373  
Hypothetical protein AGR\_C\_187 [imported] - Agrobacterium tumefaciens (strain C58, Cerec  
C:Species: Agrobacterium tumefaciens  
C>Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 11-Jan-2002  
C:Accession: B97373  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
A:Reference number: A97359; PMID:11743194  
A:Accession: B97373  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-88 <KUR>  
A:Cross-references: GB:AE007869; PIDN:AAK85939.1; PID:g15154992; GSPDB:GN00169  
C:Genetics:  
A:Gene: AGR\_C\_187  
A:Map position: circular chromosome

Query Match  
Best Local Similarity 100.0%; Score 18; DB 2; Length 88;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SUDA 4  
||||  
DB 76 SUDA 79

## RESULT 24

AB2591  
conserved hypothetical protein Atu0119 [imported] - Agrobacterium tumefaciens (strain  
C:Species: Agrobacterium tumefaciens  
C>Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002  
C:Accession: AB2591  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McCl  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam  
ster, E.W.

A:Title: The genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; PMID:11743193  
A:Accession: AB2591  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-88 <KUR>  
A:Cross-references: GB:AE008688; PIDN:AA41144.1; PID:g17738440; GSPDB:GN00186  
C:Genetics:  
A:Gene: Atu0119  
A:Map position: circular chromosome

Query Match  
Best Local Similarity 100.0%; Score 18; DB 2; Length 88;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SUDA 4  
||||  
DB 76 SUDA 79

## RESULT 25

T44780  
DNA-binding protein HU-1 [imported] - Bordetella pertussis  
C:Species: Bordetella pertussis  
C>Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 02-Sep-2000  
C:Accession: T44780  
R:Pradel, E.; Guiso, N.; Loch, C.  
submitted to the EMBL Data Library, February 1999  
A:Description: Construction and characterization of Bordetella pertussis tonb mutants  
A:Reference number: Z22838  
A:Accession: T44780  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-90 <PRA>  
A:Cross-references: EMBL:AJ132741; PIDN:CAB53382.1  
A:Experimental source: strain Tohamai  
C:Genetics:  
A:Note: huBP  
C:Superfamily: bacterial DNA-binding protein

Query Match  
Best Local Similarity 100.0%; Score 18; DB 2; Length 90;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SUDA 4  
||||  
DB 24 SUDA 27

## RESULT 26

T09285  
embryonic abundant protein EMB44 - white spruce  
C:Species: Picea glauca (white spruce)  
C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 23-Jul-1999  
C:Accession: T09285  
R:Dong, J.Z.; Dunstan, D.I.  
submitted to the EMBL Data Library, June 1996  
A:Description: Gene expression during somatic embryogenesis.  
A:Reference number: Z16588  
A:Accession: T09285  
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA  
A:Residues: 1-91 <DON>  
A:Cross-references: EMBL:U47750; NID:q1350537; PID:q1350538  
C:Genetics:  
A:Gene: EMB44  
C:Superfamily: embryonic abundant protein Em

Query Match 100.0%; Score 18; DB 2; Length 91;  
Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SLDA 4  
||||  
Db 29 SLDA 32

## RESULT 27

T09293  
hypothetical protein - white spruce  
C:Species: Picea glauca (white spruce)  
C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 23-Jul-1999  
C:Accession: T09293  
R:Donng, J.Z.; Dunstan, D.I.  
Planta 199, 459-466, 1996  
A:Title: Expression of abundant mRNAs during somatic embryogenesis of white spruce.  
A:Reference number: Z16636; MUID:96367688; PMID:8771802  
A:Accession: T09293  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-91 <DON>  
A:Cross-references: EMBL:U42464; NID:q1161168; PID:q1161169  
C:Superfamily: embryonic abundant protein Em

Query Match 100.0%; Score 18; DB 2; Length 91;  
Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SLDA 4  
||||  
Db 29 SLDA 32

## RESULT 28

E81209  
conserved hypothetical protein NMB0343 [imported] - Neisseria meningitidis (strain MC58  
C:Species: Neisseria meningitidis  
C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
C:Accession: E81209  
R:Rettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.  
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;  
Ri, H.; Qiu, H.; Vamathyan, J.; Gill, J.; Scariato, V.; Maignani, V.; Pizsa, M.  
Science 287, 1809-1815, 2000  
A:Authors: Grandl, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V.  
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
A:Reference number: A81000; MUID:20175755; PMID:10710307  
A:Accession: E81209  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-92 <TET>  
A:Cross-references: GB:AE003391; GB:AE002098; NID:g7225561; PIDN:AAF40786.1; PID:g722556  
A:Experimental source: serogroup B, strain MC58  
C:Genetics:  
A:Gene: NMB0343

Query Match 100.0%; Score 18; DB 2; Length 92;  
Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SLDA 4  
||||  
Db 60 SLDA 63

RESULT 29  
T10931  
ribosomal protein S15 - Streptomyces coelicolor

C:Species: Streptomyces coelicolor  
C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 20-Jun-2000  
C:Accession: T10931  
R:Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, August 1998  
A:Reference number: Z17215  
A:Accession: T10931  
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
A:Residues: 1-95 <PAR>  
A:Cross-references: EMBL:AL031231  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: rpsO  
C:Superfamily: Escherichia coli ribosomal protein S15; eubacterial ribosomal protein  
F.23-89/Domain: eubacterial ribosomal protein S15 homology <S15>

Query Match 100.0%; Score 18; DB 2; Length 95;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SLDA 4  
||||  
Db 2 SLDA 5

## RESULT 30

E81786  
conserved hypothetical protein NMA2144 [imported] - Neisseria meningitidis (strain Z2  
C:Species: Neisseria meningitidis  
C>Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 02-Feb-2001  
C:Accession: E81786  
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo  
: Holtroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre  
Nature 404, 502-506, 2000  
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491  
A:Reference number: A81775; MUID:20222556; PMID:10761919  
A:Accession: E81786  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-96 <PAR>  
A:Cross-references: GB:AL162758; GB:AL157959; NID:g7380672; PIDN:CAB85356.1; PID:g738  
A:Experimental source: serogroup A, strain Z2491  
C:Genetics:  
A:Gene: NMA2144

Query Match 100.0%; Score 18; DB 2; Length 96;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SLDA 4  
||||  
Db 64 SLDA 67

## RESULT 31

T12686  
hypothetical protein 63B12.12 - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C>Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 17-Nov-2000  
C:Accession: T12686  
R:Ferraz, C.; Vidal, S.; Brun, C.; Bucheton, A.; Demaille, J.G.  
submitted to the EMBL Data Library, January 1998  
A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.  
A:Reference number: Z17572  
A:Accession: T12686  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-96 <FER>  
A:Cross-references: EMBL:AL021106; NID:el371406; PID:el226203; PIDN:CA15943.1

A:Experimental source: clone cosmid 63B12  
 C:Genetics:  
 A:Cross-references: FlyBase:FBgn0000117  
 A:Introns: 27/1  
 A:Note: 63B12.12

Query Match 100.0%; Score 18; DB 2; Length 96;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
 ||||  
 DB 34 SLDA 37

## RESULT 32

JC2403  
 PMS8 homolog mismatch repair protein - human

C:Species: Homo sapiens (man)  
 C:Date: 24-Feb-1995 #sequence\_revision 24-Feb-1995 #text\_change 05-Nov-1999  
 C:Accession: JC2403  
 R:Horii, A.; Han, M.-J.; Sasaki, S.; Shimada, M.; Nakamura, Y.  
 Biochem. Biophys. Res. Commun. 204, 1257-1264, 1994  
 A:Title: Cloning, characterization and chromosomal assignment of the human genes homolog  
 A:Reference number: JC2398; MUID:95071462; PMID:7980603  
 A:Accession: JC2403  
 A:Molecule type: DNA  
 A:Residues: 1-98 <HOR>  
 A:Cross-references: DDBJ:D38440; NID:9600595; PIDN:BAA07475.1; PID:d1008055; PID:g140758  
 C:Genetics:  
 A:Gene: GDB:PMS2L6; PMS8  
 A:Cross-references: GDB:437147  
 A:Map position: 7q11.23-7q22  
 C:Keywords: DNA repair

Query Match 100.0%; Score 18; DB 2; Length 98;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
 ||||  
 DB 50 SLDA 53

## RESULT 33

S70645  
 cell division control protein CKS1 - Leishmania mexicana

N:Alternate names: protein PL2LmmCKS1; protein p13suc1 homolog  
 C:Species: Leishmania mexicana  
 C:Date: 14-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 18-Jun-1999  
 C:Accession: S70645; S57414  
 R:Mottram, J.C.; Grant, K.M.  
 Biochem. J. 316, 833-839, 1996  
 A:Title: Leishmania mexicana p12cksl, a homologue of fission yeast p13suc1, associates w  
 A:Reference number: S70645; MUID:96265049; PMID:8670159  
 A:Accession: S70645  
 A:Molecule type: DNA  
 A:Residues: 1-99 <MOT>  
 A:Cross-references: EMBL:Z49881; NID:9871408; PIDN:CAA0037.1; PID:g871409  
 C:Genetics:  
 A:Gene: CKS1  
 C:Superfamily: cell division control protein CKS1  
 C:Keywords: cell cycle control

Query Match 100.0%; Score 18; DB 2; Length 99;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
 ||||  
 DB 11 SLDA 14

RESULT 34  
 AC1127  
 hypothetical protein lmo0418 [imported] - Listeria monocytogenes (strain EGD-e)

C:Species: Listeria monocytogenes  
 C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
 C:Accession: AC1127  
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec  
 .; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl,  
 D.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkut, G.; Madueno, E.; Maitournam, A.;  
 Ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla  
 A:Title: Comparative genomics of Listeria species.  
 A:Reference number: AB1077; MUID:21537279; PMID:11679669  
 A:Accession: AC1127  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-99 <GLA>  
 A:Cross-references: GB:NC\_003210; PIDN:CAC98497.1; PID:g16409795; GSPDB:GN00177  
 A:Experimental source: strain EGD-e  
 C:Genetics:  
 A:Gene: lmo0418

Query Match 100.0%; Score 18; DB 2; Length 99;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
 ||||  
 DB 3 SLDA 6

## RESULT 35

AH1487  
 hypothetical protein lin0439 [imported] - Listeria innocua (strain C11p1262)

C:Species: Listeria innocua  
 C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
 C:Accession: AH1487  
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec  
 .; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl,  
 D.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkut, G.; Madueno, E.; Maitournam, A.;  
 Ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla  
 A:Title: Comparative genomics of Listeria species.  
 A:Reference number: AB1077; MUID:21537279; PMID:11679669  
 A:Accession: AH1487  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-99 <GLA>  
 A:Cross-references: GB:AL592022; PIDN:CAC95672.1; PID:g16412868; GSPDB:GN00178  
 A:Experimental source: strain C11p1262  
 C:Genetics:  
 A:Gene: lin0439

Query Match 100.0%; Score 18; DB 2; Length 99;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
 ||||  
 DB 3 SLDA 6

## RESULT 36

TJ35232  
 hypothetical protein SC5C7.25 SC5C7.25 - Streptomyces coelicolor

C:Species: Streptomyces coelicolor  
 C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
 C:Accession: TJ35232  
 R:Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
 submitted to the EMBL Data Library, September 1998  
 A:Reference number: Z21572

A:Accession: T35232  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-101 <SEE>  
A:Cross-references: EMBL:AL031515; PIDN:CAA20637.1; GSPDB:GN00070; SCOEDB:SC5C7.25  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCOEDB:SC5C7.25

Query Match 100.0%; Score 18; DB 2; Length 101;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
Db 22 SLDA 25

RESULT 37  
S19225  
embryonic abundant protein D-19 - upland cotton  
N:Alternate names: lea protein D-19  
C:Species: Gossypium hirsutum (upland cotton)  
C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 20-Aug-1999  
C:Accession: S19225; S04041  
R:Dure III, L.S.  
submitted to MIPS, March 1992  
A:Reference number: S19225  
A:Accession: S19225  
A:Molecule type: DNA  
A:Residues: 1-102 <BAK1>  
A:Cross-references: EMBL:X13205; NID:g18493; PIDN:CAA31593.1; PID:g18494  
R:Baker, J.; Steele, C.; Dure III, L.  
Plant Mol. Biol. 11, 277-281, 1988  
A:Title: Sequence and characterization of 6 lea proteins and their genes from cotton.  
A:Reference number: S04041  
A:Accession: S04041  
A:Molecule type: DNA  
A:Residues: 1-101, 'KINIIISHDRLATPPIYK' <BAK>  
A:Cross-references: EMBL:X13205  
A:Note: This sequence has been revised in reference S19225  
C:Genetics:  
A:introns: 49/1  
C:Superfamily: embryonic abundant protein Em

Query Match 100.0%; Score 18; DB 2; Length 102;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
Db 39 SLDA 42

RESULT 38  
T41937  
hypothetical protein U35 - human herpesvirus 7 (strain JI)  
C:Species: human herpesvirus 7  
A:Variety: strain JI  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T41937  
R:Nicholas, J.  
submitted to the EMBL Data Library, December 1995  
A:Description: Determination and analysis of the complete nucleotide sequence of human h  
A:Reference number: Z22022  
A:Accession: T41937  
A:Molecule type: DNA  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Residues: 1-104 <NIC>  
A:Cross-references: EMBL:U43400; PIDN:AAC54697.1  
A:Experimental source: strain JI  
C:Genetics:  
A:Note: U35

Query Match 100.0%; Score 18; DB 2; Length 104;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
Db 41 SLDA 44

RESULT 39  
H87712  
phosphoribosyl-ATP pyrophosphatase [imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C>Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001  
C:Accession: H87712  
R:Niernman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,  
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwin, M.L.; Haft, D.H.; Ko  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of Caulobacter crescentus.  
A:Reference number: A87249; M01D:21173698; PMID:11259647  
A:Accession: H87712  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-107 <STO>  
A:Cross-references: GB:AE005673; NID:g13425510; PIDN:AAK25700.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC3738  
C:Superfamily: phosphoribosyl-ATP pyrophosphatase hist

Query Match 100.0%; Score 18; DB 2; Length 107;  
Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
Db 83 SLDA 86

RESULT 40  
T17856  
hypothetical protein a356R - Chlorella virus PBCV-1  
C:Species: Chlorella virus PBCV-1  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T17856  
R:Graves, M.V.; Van Etten, J.L.  
submitted to the EMBL Data Library, May 1999  
A:Reference number: Z18806  
A:Accession: T17856  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-107 <GRA>  
A:Cross-references: EMBL:U42580; NID:g4028896; PIDN:AAC96724.1  
A:Experimental source: specific host Chlorella strain NC64A  
C:Genetics:  
A:Note: a356R

Query Match 100.0%; Score 18; DB 2; Length 107;  
Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
Db 57 SLDA 60

RESULT 41  
T06898  
hypothetical protein 108 - Cyanophora paradoxa cyanelle  
C:Species: cyanelle Cyanophora paradoxa  
C>Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 08-Oct-1999  
C:Accession: T06898

R:Stirewalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohner, H.J.; Bryant, D.A.  
submitted to the EMBL Data Library, July 1995  
A:Description: Nucleotide sequence of the cyanelle genome from *Cyanophora paradoxa*.  
A:Reference number: Z15840  
A:Accession: T06898  
A:Status: Preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-108 <ST>  
A:Cross-references: EMBL:U30821; NID:g1016083; PIDN:AAA81241.1; PID:g1016154  
C:Experimental source: strain Pringsheim LB555  
C:Genetics:  
A:Gene: ori108  
A:Genome: cyanelle  
C:Keywords: cyanelle

Query Match 100.0%; Score 18; DB 2; Length 108;  
Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SUDA 4  
|||  
Db 40 SUDA 43

## RESULT 42

hypothetical protein SCL1A9.28c SCL1A9.28c - Streptomyces coelicolor

C:Species: Streptomyces coelicolor  
C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
C:Accession: T34677  
R:Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.  
submitted to the EMBL Data Library, December 1998  
A:Reference number: Z21552  
A:Accession: T34677  
A:Status: Preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-108 <SAU>  
A:Cross-references: EMBL:AL034446; PIDN:CAA22398.1; GSPDB:GN00070; SCORDB:SCL1A9.28c  
C:Genetics:  
A:Gene: SCORDB:SCL1A9.28c

Query Match 100.0%; Score 18; DB 2; Length 108;  
Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SUDA 4  
|||  
Db 2 SUDA 5

## RESULT 43

hypothetical protein APE0568 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jun-2000  
C:Accession: H72641  
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudon, Y.; Yamazaki, J.; KDNA Res. 6, 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, *Aeropyrum*  
A:Reference number: A72450; MUID:99310339; PMID:10382966  
A:Accession: H72641  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-109 <KAM>  
A:Cross-references: DDBJ:AP000060; NID:95104188; PIDN:BAA79536.1; PID:J1043322; PID:9510  
C:Genetics:  
A:Experimental source: strain K1  
A:Gene: APE0568  
C:Superfamily: Aeropyrum pernix hypothetical protein APE0568

Query Match 100.0%; Score 18; DB 2; Length 109;

Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SUDA 4  
|||  
Db 25 SUDA 28

## RESULT 44

sigma regulatory factor - Chlamydia pneumoniae (strain CWL029)

C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 05-May-2000  
C:Accession: D72019  
R:Kaiman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, Nature Genet. 21, 385-389, 1999  
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.  
A:Reference number: A72000; MUID:99206606; PMID:10192388  
A:Accession: D72019  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-110 <ARN>  
A:Cross-references: GB:AE001671; GB:AE001363; NID:g4377226; PIDN:AAD19047.1; PID:g437  
C:Genetics:  
A:Gene: rsbv\_2  
C:Superfamily: sporulation protein stage II

Query Match 100.0%; Score 18; DB 2; Length 110;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SUDA 4  
|||  
Db 19 SUDA 22

## RESULT 45

sigma regulatory factor [imported] - Chlamydia pneumoniae (strain J138)

C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 17-May-2002  
C:Accession: C86604  
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Nucleic Acids Res. 28, 2311-2314, 2000  
A:Title: Comparison of whole genome sequences of Chlamydia pneumoniae J138.  
A:Reference number: A86491; MUID:20330349; PMID:10871362  
A:Accession: C86604  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-110 <STO>  
A:Cross-references: GB:BA000008; NID:g8979283; PIDN:BAA99117.1; GSPDB:GN00142  
C:Genetics:  
A:Experimental source: strain J138  
A:Gene: rsbv\_2  
C:Superfamily: sporulation protein stage II

Query Match 100.0%; Score 18; DB 2; Length 110;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SUDA 4  
|||  
Db 19 SUDA 22

Search completed: February 6, 2003, 11:23:08  
Job time : 12.5 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: February 6, 2003, 11:16:00 : Search time 5.3333 seconds  
(without alignments)  
31.107 Million cell updates/sec

Title: PAT943-4  
Perfect score: 18  
Sequence: 1 slda 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 1064

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 100%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	100.0	40	1	HPT_RABIT
2	18	100.0	62	1	DMP1_PSESP
3	18	100.0	62	1	XYH_PSEPU
4	18	100.0	75	1	EXTS_CHUPN
5	18	100.0	79	1	GP56_BPSP1
6	18	100.0	87	1	YEB2_ARCFU
7	18	100.0	85	1	Y04A_BP74
8	18	100.0	91	1	EML_PICGL
9	18	100.0	95	1	RS15_STRCO
10	18	100.0	96	1	ES98_DROME
11	18	100.0	99	1	CKS1_LETME
12	18	100.0	101	1	PHS_STRCO
13	18	100.0	102	1	LE19_GOSHI
14	18	100.0	107	1	HIS2_CAVCR
15	18	100.0	108	1	YC50_CVAPA
16	18	100.0	112	1	ACPW_SCHPO
17	18	100.0	115	1	SFP3_BOVIN
18	18	100.0	116	1	Y998_RHIME
19	18	100.0	118	1	YEO6_HAEIN
20	18	100.0	120	1	CHH4_PEMMO
21	18	100.0	122	1	RL22_HELPU
22	18	100.0	122	1	RL22_HELPU
23	18	100.0	124	1	RS12_XYLF
24	18	100.0	126	1	THN1_WHEAT
25	18	100.0	130	1	UL33_HSV11
26	18	100.0	133	1	YY08_MYCTU
27	18	100.0	134	1	SFP1_BOVIN
28	18	100.0	134	1	YH11_MYCLE
29	18	100.0	135	1	YNP0_YEAST
30	18	100.0	136	1	THN2_WHEAT
31	18	100.0	136	1	THNB_WHEAT
32	18	100.0	137	1	NXT1_CAEEL
33	18	100.0	138	1	NUSB_YERPE

34	18	100.0	144	1	PSB2_ECOLI
35	18	100.0	147	1	HFAA_CADCR
36	18	100.0	147	1	RS12_SULSO
37	18	100.0	147	1	RS12_SULTO
38	18	100.0	150	1	TPCB_HOMAM
39	18	100.0	151	1	MPRA2_PSEAE
40	18	100.0	151	1	RS12_SULAC
41	18	100.0	151	1	Y4JR_RHISN
42	18	100.0	154	1	COBE_PSEDE
43	18	100.0	156	1	VATL_HELVI
44	18	100.0	157	1	U51_HCMVA
45	18	100.0	160	1	RISB_MYCTU

## ALIGNMENTS

RESULT 1	HPT_RABIT	STANDARD:	PRT:	40 AA.
AC	P19007;			
DT	01-NOV-1990 (Rel. 16, Created)			
DT	01-NOV-1990 (Rel. 16, Last sequence update)			
DE	01-OCT-1996 (Rel. 34, Last annotation update)			
DE	Haptoglobin beta chain (Fragment).			
GN	HP.			
OS	Oryctolagus cuniculus (Rabbit).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.			
OX	NCBI_Taxid=9986;			
RN	[1]			
RP	SEQUENCE.			
RX	MEDLINE=77025019; PubMed=975782;			
RA	Kurosky A., Kim H.H., Touchstone B.;			
RT	"Comparative sequence analysis of the N-terminal region of rat,			
RT	rabbit, and dog haptoglobin beta-chains.";			
RL	Comp. Biochem. Physiol. 55B:453-459(1976).			
CC	-1- FUNCTION: HAPTOGLOBIN COMBINES WITH FREE PLASMA HEMOGLOBIN,			
CC	PREVENTING LOSS OF IRON THROUGH THE KIDNEYS AND PROTECTING THE			
CC	KIDNEYS FROM DAMAGE BY HEMOGLOBIN, WHILE MAKING THE HEMOGLOBIN			
CC	ACCESSIBLE TO DEGRADATIVE ENZYMES.			
CC	-1- SUBUNIT: TETRAMER OF TWO ALPHA AND TWO BETA CHAINS.			
CC	-1- TISSUE SPECIFICITY: SYNTHESIZED IN THE LIVER, THEN RELEASED			
CC	IN PLASMA.			
CC	-1- SIMILARITY: THE BETA CHAIN IS CLEARLY RELATED TO SERINE PROTEASES,			
CC	BUT HAPTOGLOBIN HAS NO PROTEOLYTIC ACTIVITY. PROBABLY BECAUSE THE			
CC	POSITIONS CORRESPONDING TO THE PROTEOLYTIC ACTIVE SITE RESIDUES IN			
CC	PROTEASES ARE OCCUPIED BY DIFFERENT AMINO ACIDS IN HAPTOGLOBIN.			
CC	PIR: A26503; A26503.			
DR	InterPro: IPR001254; Ser-protease_Try.			
DR	Pfam: PF00089; trypsin_1.			
DR	PROSITE: PS50240; TRYPSIN_DOM; PARTIAL.			
KW	Serine protease homolog; Plasma; Hemoglobin-binding; Liver.			
FT	NON_TER 40			
SO	SEQUENCE 40 AA; 4299 MW; 215B2003101B7B84 CRC64;			
Query Match				
Best Local Similarity 100.0%; Score 18; DB 1; Length 40;				
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1 SLD4 4			
DB	5 SLD4 8			
RESULT 2				
ID	DMP1_PSESP	STANDARD:	PRT:	62 AA.
AC	P49172;			
DT	01-FEB-1996 (Rel. 33, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			

```

DE 4-oxalocrotonate tautomerase (EC 5.3.2.-) (4-OT).
GN DMPI.
OS Pseudomonas sp. (strain CF600).
OC Plasmid pV1150.
CC Bacteria; Proteobacteria.
OX NCBI_TaxID=306;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92121108; PubMed=1732207;
RA Shingler V., Marklund U., Fowlowski J.;
RT "Nucleotide sequence and functional analysis of the complete
RT phenol/3,4-dimethylphenol catabolic pathway of Pseudomonas sp.
RT strain CF600."
RL J. Bacteriol. 174:711-724(1992).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=96146412; PubMed=8547259;
RA Subramanya H.S., Roper D.I., Dauter Z., Dodson E.J., Davies G.J.,
RA Wilson K.S., Wigley D.B.;
RT "Enzymatic ketonization of 2-hydroxymuconate: specificity and
RT mechanism investigated by the crystal structures of two isomerases."
RL Biochemistry 35:792-802(1996).
CC -1- FUNCTION: CATALYZES THE KETONIZATION OF 2-HYDROXYMUCONATE
CC STEREOSELECTIVELY TO YIELD 2-OXO-3-HEXENEDIOLATE.
CC -1- PATHWAY: META-CLEAVAGE PATHWAY FOR THE DEGRADATION OF TOLUENE,
CC M-XYLENE AND P-XYLENE.
CC -1- SUBUNIT: HOMOHETEROMER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X60835; CAA43229.1; -.
DR PDB: 1OTF; 03-APR-96.
DR InterPro: IPR004370; Taut.
DR Pfam: PF01361; Tautomerase; 1.
DR ProDom: PD019232; Tautomerase; 1.
DR TIGRFAMs: TIGR00013; taut; 1.
DR TrEMBL: X60835; CAA43229.1; Tautomerase; 1.
KW isomerase; Plasmid; Aromatic hydrocarbons catabolism; 3D-structure.
FT INIT_MET 0
FT ACT_SITE 1
FT SEQUENCE 62 AA; 6974 MW; CEDAB0P92D1E1E0F CRC64;
SQ
Query Match Best Local Similarity 100.0%; Score 18; DB 1; Length 62;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLDA 4
Db 30 SLDA 33

```

## RESULT 3

```

XYLH_PSEPU STANDARD; PRT: 62 AA.
AC Q01468;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE 4-oxalocrotonate tautomerase (EC 5.3.2.-) (4-OT).
GN XYLH.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-33.
RC STRAIN=mt-2;

```

```

RX MEDLINE=92388122; PubMed=1339435;
RA Chen L.H., Kenyon G.L., Curtin F., Harayama S., Benbenek M.E.,
RA Hajipour G., Whitman C.P.;
RT "4-oxalocrotonate tautomerase, an enzyme composed of 62 amino acid
RT residues per monomer."
RL J. Biol. Chem. 267:17716-17721(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=mt-2;
RC MEDLINE=93288011; PubMed=8510667;
RA Harayama S., Reik M.;
RT "Comparison of the nucleotide sequences of the meta-cleavage pathway
RT genes of TOL plasmid pMT0 from Pseudomonas putida with other meta-
RT cleavage genes suggests that both single and multiple nucleotide
RT substitutions contribute to enzyme evolution."
RL Mol. Gen. Genet. 239:81-89(1993).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=96146412; PubMed=8547259;
RA Subramanya H.S., Roper D.I., Dauter Z., Dodson E.J., Davies G.J.,
RA Wilson K.S., Wigley D.B.;
RT "Enzymatic ketonization of 2-hydroxymuconate: specificity and
RT mechanism investigated by the crystal structures of two isomerases."
RL Biochemistry 35:792-802(1996).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RC STRAIN=mt-2;
RX MEDLINE=98453316; PubMed=9778344;
RA Taylor A.B., Czerwinski R.M., Johnson W.H. Jr., Whitman C.P.,
RA Hackert M.L.;
RT "Crystal structure of 4-oxalocrotonate tautomerase inactivated by 2-
RT oxo-3-pentynoate at 2.4-A resolution: analysis and implications for
RT the mechanism of inactivation and catalysis."
RL Biochemistry 37:14692-14700(1998).
RN [5]
RP STRUCTURE BY NMR, AND ACTIVE SITE.
RX MEDLINE=96146413; PubMed=8547260;
RA Silvers J.T., Abeygunawardana C., Mildvan A.S., Hajipour G.,
RA Whitman C.P., Chen L.H.;
RT "Catalytic role of the amino-terminal proline in 4-oxalocrotonate
RT tautomerase: affinity labeling and heteronuclear NMR studies."
RL Biochemistry 35:803-813(1996).
CC -1- FUNCTION: CATALYZES THE KETONIZATION OF 2-HYDROXYMUCONATE
CC STEREOSELECTIVELY TO YIELD 2-OXO-3-HEXENEDIOLATE.
CC -1- PATHWAY: META-CLEAVAGE PATHWAY FOR THE DEGRADATION OF TOLUENE,
CC M-XYLENE AND P-XYLENE.
CC -1- SUBUNIT: HOMOHETEROMER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M95650; AAA26046.1; -.
DR EMBL: M94186; AAA25694.1; -.
DR PIR: A43397; A43397.
DR PIR: S35225; S35225.
DR PDB: 1AUF; 13-JAN-99.
DR InterPro: IPR004370; Taut.
DR Pfam: PF01361; Tautomerase; 1.
DR ProDom: PD019232; Tautomerase; 1.
DR TIGRFAMs: TIGR00013; taut; 1.
KW isomerase; Plasmid; Aromatic hydrocarbons catabolism; 3D-structure.
FT INIT_MET 0
FT ACT_SITE 1
FT SEQUENCE 62 AA; 6811 MW; 23E6DC494A12254F CRC64;
SQ
Query Match Best Local Similarity 100.0%; Score 18; DB 1; Length 62;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



OY 1 SLDA 4  
1111  
DB 30 SLDA 33

## RESULT 4

EXTS.CHLPN STANDARD: PRT: 75 AA.  
ID EX75.CHLPN  
AC 09K1Y4;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Probable exodeoxyribonuclease VII small subunit (EC 3.1.11.6)  
DE (Exonuclease VII small subunit).  
GN XSEB OR CPN1061.1 OR CP0788.  
OS Chlamydia pneumoniae (Chlamydia pneumoniae).  
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia. *Chlamydia*.  
OX NCBI\_TaxID=83558;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CWL029;  
RX MEDLINE=99206606; PubMed=10192388;  
RA Kaiman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,  
RT Olinger L., Grimwood J., Davis R.W., Stephens R.S.;  
"Comparative genomes of Chlamydia pneumoniae and C. trachomatis.",  
Nat. Genet. 21:385-389(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-AR39;  
RX MEDLINE=20150255; PubMed=10684935;  
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,  
RT White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,  
RA Liner K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,  
RA Gwin M., Nelson W., DeBo R., Kolonay J., McClarty G., Salzberg S.L.,  
RT Eisen J., Fraser C.M.;  
"Genome sequences of Chlamydia trachomatis Mopn and Chlamydia  
RT pneumoniae AR39.",  
Nucleic Acids Res. 28:1397-1406(2000).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-J138;  
RX MEDLINE=20330349; PubMed=10871362;  
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,  
RT Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;  
"Comparison of whole genome sequences of Chlamydia pneumoniae J138  
RT from Japan and CWL029 from USA.",  
Nucleic Acids Res. 28:2311-2314(2000).  
CC -1- FUNCTION: BIDIRECTIONALLY DEGRADES SINGLE-STRANDED DNA INTO LARGE  
CC ACID-INSOLUBLE OLIGONUCLEOTIDES, WHICH ARE THEN DEGRADED FURTHER  
CC INTO SMALL ACID-SOLUBLE OLIGONUCLEOTIDES (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5'- to 3'-  
CC or 3'- to 5'-direction to yield nucleoside 5'-phosphates.  
CC -1- SUBUNIT: HETEROOLIGOMER COMPOSED OF LARGE AND SMALL SUBUNITS (BY  
CC SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -1- SIMILARITY: BELONGS TO THE XSEB FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: AE001686; -; NOT\_ANNOTATED\_CDS.  
CC EMBL: AE002238; AAF38587.1; -;  
CC EMBL: AP002548; -; NOT\_ANNOTATED\_CDS.  
CC TIGR: CP0788; -;  
CC InterPro: IPR003761; Exonuc\_VII\_S.  
CC Pfam: PF02609; Exonuc\_VII\_S; 1.  
CC TIGRPFAMs: TIGR01280; xseb; 1.

KW Hydrolase; Nuclease; Exonuclease; Complete proteome.  
SQ SEQUENCE 75 AA; 8809 MW; AA05874AF83728 CRC64;

Query Match 100.0%; Score 18; DB 1; Length 75;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
1111  
DB 26 SLDA 29

## RESULT 5

GP56\_BPSP1 STANDARD: PRT: 79 AA.  
ID GP56\_BPSP1  
AC 048410;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Putative gene protein 56.  
GN 56.  
OS Bacteriophage SP01.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;  
OC SP01-like viruses.  
OX NCBI\_TaxID=10685;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98327781; PubMed=9657951;  
RA Stewart C.R., Gaslightwala I., Hinata K., Krolkowski K.A.,  
RT Needleman D.S., Peng A.S.-Y., Peterman M.A., Tobias A., Wei P.;  
"Genes and regulatory sites of the 'host-takeover module' in the  
RT terminal redundancy of Bacillus subtilis bacteriophage SP01.",  
Virology 246:329-340(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: AF031901; AAC29025.1; -;  
CC DR Hypothetical protein.  
KW SEQUENCE 79 AA; 9305 MW; 9783ABB51E0A609A CRC64;  
SQ  
Query Match 100.0%; Score 18; DB 1; Length 79;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## RESULT 6

YE82\_ARCFU STANDARD: PRT: 85 AA.  
ID YE82\_ARCFU  
AC 028790;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein AF1482.  
GN AF1482.  
OS Archaeoglobus fulgidus.  
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;  
OC Archaeoglobaceae; Archaeoglobus.  
OX NCBI\_TaxID=2234;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;  
RX MEDLINE=98049343; PubMed=9389475;  
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,

```

RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyprides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., Mckenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodok A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001000; AAB89770.1; -.
DR TIGR; AF1482; -.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 12 34 POTENTIAL.
FT TRANSSEM 49 71
SQ SEQUENCE 85 AA; 9864 MW; 1E2C08D875699B65 CRC64;

Query Match
Best Local Similarity 100.0%; Score 18; DB 1; Length 85;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4
   ||||
Db 24 SLDA 27

RESULT 7
Y04A_BP74 STANDARD; PRT; 87 AA.
ID Y04A_BP74
AC P07083;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical 9.8 kDa protein in Gp55-nrdg intergenic region.
GN Y04A OR 55.1.
OS Bacteriophage T4.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC T4-like viruses.
OX NCBI_TaxID=10665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C;
RX MEDLINE=87203398; PubMed=3575111;
RA Tomaszewski J., Rueger W.;
RT "Nucleotide sequence and primary structures of gene products coded
RT for by the T4 genome between map positions 48.266 kb and 39.166 kb.";
RL Nucleic Acids Res. 15:3632-3633(1987).
RN [2]
RP SEQUENCE FROM N.A.
RA Kuter E., Arita A., Kunisawa T., Tsugita A., Mosig G.,
RA Mesyanzhinov V., Ruger W., Stidham T., Thomas E.;
RT "Bacteriophage T4 genome analysis.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL; Y00122; CAA68320.1; -.
DR EMBL; AF158101; AAD42492.1; -.
DR PIR; F30292; ZGBPT9.
KW Hypothetical protein.
SQ SEQUENCE 87 AA; 9845 MW; 9962EF2096CC419 CRC64;

Query Match
Best Local Similarity 100.0%; Score 18; DB 1; Length 87;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4
   ||||
Db 58 SLDA 61

RESULT 8
EMBL_PICGL
ID EML_PICGL STANDARD; PRT; 91 AA.
AC Q40864;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Em-like protein.
OS Picea glauca (White spruce).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
OX NCBI_TaxID=3330;
RN [1]
RP SEQUENCE FROM N.A.
RA Dong J.Z., Dunstan D.I.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; I42464; AAA85366.1; -.
DR InterPro; IPR000389; Seed protein.
DR Pfam; PF00477; Seed_protein; 1.
DR PROSITE; PS00431; SMALL_HYDR_PLANT_SEED; 1.
KW Seed; Seed embryo.
SQ SEQUENCE 91 AA; 9771 MW; 8FD9D4F41E0A195E CRC64;

Query Match
Best Local Similarity 100.0%; Score 18; DB 1; Length 91;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4
   ||||
Db 29 SLDA 32

RESULT 9
R515_STRCO
ID R515_STRCO STANDARD; PRT; 95 AA.
AC O86655;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S15.
GN RPSO OR SC05736 OR SC3C3.22.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1902;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-A3(2) / M145;
RX MEDLINE-21996410; PubMed-12000953;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kisser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.M., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Krieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinovitch E., Rajadream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
*Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE 16S RIBOSOMAL RNA BINDING
CC PROTEINS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE S15P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AL031231; CA20271.1;
CC HSP: P80378; IAB3.
CC InterPro: IPR005290; RS15_bact.
CC Pfam: PF00312; Ribosomal_S15_1.
CC TIGRfam: TIGR00952; S15_bact_1.
CC PROSITE: PS00362; RIBOSOMAL_S15; 1.
CC RiboSOMeL protein: RNA-binding; Complete proteome.
SQ SEQUENCE 95 AA; 10790 MW; A5E3A8E04CD8182 CRC64;

Query Match 100.0%; Score 18; DB 1; Length 95;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SIDA 4
Db 2 SIDA 5

RESULT 10
ES98_DROME STANDARD; PRT; 96 AA.
AC 097420;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE EST00098 protein homolog.
GN EG:63B12.12 OR CG14818.
OS Drosophila melanogaster (Fruit fly).
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
ON NCBI_TaxID=7227;
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Cealniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Bailley R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

```

```

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dubin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostali M., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
RA Jaisli D., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Modarri C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclob J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Ventler E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Ventler J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN-Oregon-R.
RX MEDLINE-20196011; PubMed-10731137;
RA Benos P.V., Galt M.K., Ashburner M., Murphy L., Harris D.,
RA Barrell B.G., Ferraz C., Vidal S., Brun C., Demallies J., Cadieu E.,
RA Dreano S., Gloux S., Leleure V., Mottier S., Galbert F., Borkova D.,
RA Minano B., Kafatos F.C., Louis C., Siden-Kiamos I., Bolshakov S.,
RA Papagiannakis G., Spanos L., Cox S., Madueno E., de Pablos B.,
RA Modolell J., Peter A., Schoettler P., Werner M., Mourikioti F.,
RA Beinert N., Dowe G., Scheifer U., Jaekle H., Bucheton A.,
RA Callister D.M., Campbell L.A., Darlamsou A., Henderson N.S.,
RA McMillan P.J., Salles C., Tait E.A., Valenti P., Saunders R.D.C.,
RA Glover D.M.;
RT "From sequence to chromosome: the tip of the X chromosome of D.
RT melanogaster";
RL Science 287:2220-2222(2000).
CC -1- SIMILARITY: BELONGS TO THE UPF0184 (EST00098) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE003422; AAF45682.1;
CC EMBL: AL021106; CA15943.1;
CC FlyBase: FBgn026088; EG:63B12.12.
CC InterPro: IPR005374; UPF0184.
CC Pfam: PF03670; UPF0184; 1.
CC KW Hypothetical protein; Colled coll.
CC DOMAIN 21 77 COLLED COLL. (POTENTIAL).
FT SEQUENCE 96 AA; 10404 MW; 98452A2233F31BF CRC64;
SO

Query Match 100.0%; Score 18; DB 1; Length 96;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SIDA 4
Db 34 SIDA 37

```

```

RESULT 11
CKSI_LEIME          STANDARD;          PRT:          99 AA.
ID   CKSI_LEIME
AC   Q25330;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DE   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Cyclin-dependent kinases regulatory subunit (P12LMCKS1).
OS   Leishmania mexicana.
OC   Eukaryota; Eukaryotes; Kinetoplastida; Trypanosomatidae; Leishmania.
OX   NCBI_TaxID=5665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=M379 / MNYC / B2 / 62;
RX   MEDLINE=96265049; PubMed=8670159;
RA   Mottram J.C., Grant K.M.;
RT   "Leishmania mexicana p12cksl, a homologue of fission yeast p13suc1,
RL   Biochem. J. 316:833-839(1996).
CC   -1- FUNCTION: BINDS TO THE CATALYTIC SUBUNIT OF THE CYCLIN DEPENDENT
CC   KINASES (CDC2) AND IS ESSENTIAL FOR THEIR BIOLOGICAL FUNCTION.
CC   -1- SUBUNIT: FORMS AN HOMODIMER THAT CAN PROBABLY BIND SIX KINASE
CC   SUBUNITS (BY SIMILARITY).
CC   -1- SIMILARITY: BELONGS TO THE CKS FAMILY.
-----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@sib-sib.ch).
-----
DR   EMBL; Z49881; CAA90037.1; -
DR   HSSP; P33551; 1BUH.
DR   InterPro: IPR000789; Cyc_dep_kin_rsub.
DR   Pfam: PF01111; CKS_1.
DR   PRINTS; PR00296; CYCLINKINASE.
DR   ProDom; PD005152; Cyc_dep_kin_rsub; 1.
DR   PROSITE; PS00944; CKS_1; 1.
DR   PROSITE; PS00945; CKS_2; 1.
KW   Cell division.
SQ   SEQUENCE 99 AA; 11832 MW; 5F2C766A9AD5222 CRC64;

Query Match          100.0%; Score 18; DB 1; Length 99;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY   1 SLDA 4
DB   11 SLDA 14

RESULT 12
PHS_STRCO
AC   086722;          STANDARD;          PRT:          101 AA.
ID   PHS_STRCO
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Putative pterin-4-alpha-carbinolamine dehydratase (EC 4.2.1.96) (PHS)
DE   (4-alpha-hydroxy-tetrahydropterin dehydratase) (Pterin carbinolamine
DE   dehydratase) (PCD).
GN   SC06540 OR SC5C7.25.
OS   Streptomyces coelicolor.
OC   Bacteria; Actinobacteria; Actinobacteridae; Streptomyces.
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2) / M145;
RX   MEDLINE=2196410; PubMed=12000953;

```

```

RA   Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA   Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA   Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA   Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA   Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA   Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA   Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA   Warren T., Wietzorrek A., Woodward J., Barrall B.G., Parkhill J.,
RA   Hopwood D.A.;
RT   "Complete genome sequence of the model actinomycete Streptomyces
RT   coelicolor A3(2).";
RL   Nature 417:141-147(2002).
CC   -1- CATALYTIC ACTIVITY: (6R)-6-(L-erythro-1,2-dihydroxypropyl)-
CC   5,6,7,8-tetrahydro-4a-hydroxypterin = (6R)-6-(L-erythro-1,2-
CC   dihydroxypropyl)-7,8-dihydro-6H-pterin + H(2)O.
CC   -1- SIMILARITY: BELONGS TO THE PTERIN-4-ALPHA-CARBINOAMINE
CC   DEHYDRATASE FAMILY.
-----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@sib-sib.ch).
-----
DR   EMBL; AL031515; CAA20637.1; -
DR   HSSP; P80095; 1DCO.
DR   InterPro: IPR001533; Trans_pterindh.
DR   Pfam; PF01329; Pterin_4a; 1.
DR   ProDom; PD007262; Trans_pterindh; 1.
KW   Hypothetical protein; Lyase; Complete proteome.
SQ   SEQUENCE 101 AA; 10822 MW; DB175F37A11C159 CRC64;

Query Match          100.0%; Score 18; DB 1; Length 101;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY   1 SLDA 4
DB   22 SLDA 25

RESULT 13
LE19_GOSHI
AC   P09443;          STANDARD;          PRT:          102 AA.
ID   LE19_GOSHI
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   Late embryogenesis abundant protein p-19 (LEA D-19).
OS   Gossypium hirsutum (Upland cotton).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Malvales; Malvaceae; Gossypium.
OX   NCBI_TaxID=3635;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Coker 201; TISSUE=Seed;
RA   Baker J., Steele C., Dure L. III;
RT   "Sequence and characterization of 6 Lea proteins and their genes from
RT   cotton.";
RL   Plant Mol. Biol. 11:277-291(1988).
RN   [2]
RP   REVISIONS.
RA   Dure L. III;
RL   Submitted (JAN-1992) to the SWISS-PROT data bank.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Galau G.A., Wang H.Y., Hughes D.W.;
RL   Submitted (XXX-1991) to the EMBL/GenBank/DBJ databases.
CC   -1- FUNCTION: LEA PROTEINS ARE LATE EMBRYONIC PROTEINS ABUNDANT IN
CC   HIGHER PLANT SEED EMBRYOS. THERE ARE TWO SUBSETS OF LEA PROTEINS

```

(5A, AND 5B), THE FIRST ONES ARE EXPRESSED WHEN THE COTYLEDON WEIGHT REACH 80 MG AND THE SECOND SET ARE EXPRESSED ABOVE 100 MG. THE FUNCTION OF THOSE PROTEINS IS NOT KNOWN.

-1- INDUCTION: BY ABSCISIC ACID (ABA).

-1- MISCELLANEOUS: THIS IS A SEP 5B PROTEIN.

-1- SIMILARITY: BELONGS TO THE SMALL HYDROPHILIC PLANT SEED PROTEIN FAMILY.

-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL: M19387; AAA92729.1; ALT\_SEQ.  
DR EMBL: X13205; CAA31593.1; -  
DR EMBL: M73751; AAA33062.1; -  
DR EMBL: M73752; AAA33060.1; -  
DR PIR: S19225; S19225.  
DR InterPro: IPR000389; Seed\_protein.  
DR Pfam: PF00477; seed\_protein; 1.  
DR ProSite: PS00431; SMALL\_HYDR\_PLANT\_SEED; 1.  
KW Seed: Seed embryo.  
SQ SEQUENCE 102 AA; 11072 MW; FCC74BEF167ECD27 CRC64;

Query Match 100.0%; Score 18; DB 1; Length 102;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
DB 39 SLDA 42

-----

RESULT 14  
HIS2\_CAUCR STANDARD; PRT; 107 AA.

AC Q9A228; HIS2\_CAUCR  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Phosphoribosyl-ATP pyrophosphatase (EC 3.6.1.31) (PRA-PH).  
GN HIS2 OR CC3738.  
OS Caulobacter crescentus.  
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
OC Caulobacter.  
OX NCBI\_TaxID=155892;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 19089 / CB15.  
RX MEDLINE=21173698; PubMed=11259647;  
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Madacke J.R.,  
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Hatt D.H.,  
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty K.,  
RA Osterback T., Tran K., Wolf A., Yamathavan J., Ermolaeva M., White O.,  
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
RA "Complete genome sequence of Caulobacter crescentus";  
RA Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
RL  
CC -1- CATALYTIC ACTIVITY: 1-(5-phosphoribosyl)-ATP + H(2)O = 1-(5-phosphoribosyl)-AMP + diphosphate.  
CC -1- PATHWAY: Histidine biosynthesis; second step.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -1- SIMILARITY: BELONGS TO THE PRA-PH FAMILY.

-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL: AE006031; AAK25700.1; -  
DR TIGR: CC3738; -  
DR InterPro: IPR002497; PRA-PH.  
DR Pfam: PF01503; PRA-PH; 1.  
DR ProDom: PD002611; PRA-PH; 1.  
KW Histidine biosynthesis; Hydrolase; Complete proteome.  
SQ SEQUENCE 107 AA; 11037 MW; 93280E1FAA8629ED CRC64;

Query Match 100.0%; Score 18; DB 1; Length 107;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
DB 83 SLDA 86

-----

RESULT 15  
YC50\_CYAPA STANDARD; PRT; 108 AA.

AC P48326;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical 12.7 kDa protein ycf50 (ORF108).  
GN YCF50.  
OS Cyanophora paradoxa.  
OC Cyanophora.  
OC Eukaryota; Glaucocystophyceae; Cyanophoraceae; Cyanophora.  
OX NCBI\_TaxID=2762;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LB555 / Pringsheim;  
RA Stirewalt V.L., Michalowski C.B., Loeffelhardt W., Bohnert H.J.,  
RA Bryant D.A.;  
RT "Nucleotide sequence of the cyanelle DNA from Cyanophora paradoxa";  
RT Plant Mol. Biol. Rep. 13:327-332(1995).  
RL  
RN  
RP SEQUENCE FROM N.A.  
RC STRAIN=LB555 / Pringsheim;  
RA Loeffelhardt W., Stirewalt V.L., Michalowski C.B., Annarella M.,  
RA Parley J.Y., Schluchter W.M., Chung S., Newmann-Spallart C.,  
RA Steiner J.K., Jakowitsch J., Bohnert H.J., Bryant D.A.;  
RT "The complete complexity of the cyanelle genome of Cyanophora paradoxa";  
RT the genetic complexity of a primitive plastid";  
RT (in) Schenk H.E.A., Herrmann R., Jeon K.W., Mueller N.E.,  
RL Schwemmler W. (eds.);  
RL Eukaryotism and Symbiosis, pp.40-48, Springer-Verlag, Heidelberg  
RL (1997).  
CC -1- SIMILARITY: BELONGS TO THE YCF50 FAMILY.

-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL: U30821; AAA81241.1; -  
KW Hypothetical protein; Cyanelle.  
SQ SEQUENCE 108 AA; 12698 MW; 13AD8C826B4B58CB CRC64;

Query Match 100.0%; Score 18; DB 1; Length 108;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
DB 40 SLDA 43

```

RESULT 16
ACPM_SCHPO ID ACPM_SCHPO STANDARD: PRT: 112 AA.
AC 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative acyl carrier protein, mitochondrial precursor (ACP) (NADH-
DE ubiquinone oxidoreductase 9.6 kDa subunit).
DE SPAC4H3.09.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
NCBI_TaxID=4896;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Jones M., Leather S., McDonald S., McLean J.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkhardt G., Aert R., Robben J., Gymnopoulos B.,
RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Mambutt R., Punnelle B.,
RA Goffeau A., Cadieu E., Diano S., Gloux S., Lelaure V., Motlier S.,
RA Galibert F., Aves S.V., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Rhode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shapkovskii G.V., Ussery D., Barrell B.G., Nurse P.;
RA "The genome sequence of Schizosaccharomyces pombe."
RT Nature 415:871-880(2002).
RL
CC -1- FUNCTION: IT IS POSSIBLE THAT THIS ACP IS INVOLVED IN THE
CC SYNTHESIS OF VERY-LONG-CHAIN FATTY ACIDS (BY SIMILARITY).
CC -1- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 30 DIFFERENT SUBUNITS.
CC -1- SUBCELLULAR LOCATION: Mitochondrial (By similarity).
CC -1- SIMILARITY: CONTAINS 1 ACYL CARRIER DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: Z69380; CAA93348.1; -.
CC HSSP: P02901; IACP.
CC InterPro: IPR003331; Acyl_carrier.
CC InterPro: IPR003880; Pantne_attach.
CC Pfam: PF00550; pp-binding; 1.
CC ProDom: PD000887; Acyl_carrier; 1.
CC TIGRfam: TIGR00517; acyl_carrier; 1.
CC PROSITE: PS00012; PHOSPHOPANTHETHEINE; 1.
CC PROSITE: PS50075; ACP DOMAIN; 1.
CC Hypothetical protein: Fatty acid biosynthesis; Phosphopantetheine;
CC Mitochondrion; Transit peptide; Oxidoreductase.
CC TRANSIT 1 28 MITOCHONDRION (POTENTIAL).

```

```

FT CHAIN 29 112 PUTATIVE ACYL CARRIER PROTEIN.
FT BINDING 69 PHOSPHOPANTHETHEINE (POTENTIAL).
SQ SEQUENCE 112 AA; 12519 MW; 969C777FF3622FE0D CRC64;
Query Match
Best Local Similarity 100.0%; Score 18; DB 1; Length 112;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLDA 4
Db 69 SLDA 72
RESULT 17
SEFP3_BOVIN ID SEFP3_BOVIN STANDARD: PRT: 115 AA.
AC P04557;
DE 13-AUG-1987 (Rel. 05, Created)
DE 13-AUG-1987 (Rel. 05, Last sequence update)
DE 01-OCT-1996 (Rel. 34, Last annotation update)
DE Seminal plasma protein BSP-A3.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovine; Bos.
NCBI_TaxID=9913;
[1]
RN SEQUENCE.
RX MEDLINE=87270621; PubMed=3606570;
RA Seidah N.G., Manjunath P., Rochmont J., Sairam M.R., Chretien M.,
RT "Complete amino acid sequence of BSP-A3 from bovine seminal plasma.
RT Homology to PDC-109 and to the collagen-binding domain of
RT fibronectin."
RL Biochem. J. 243:195-203(1987).
CC -1- FUNCTION: THE BSP-A PROTEINS FROM SEMINAL PLASMA EXHIBIT BOTH
CC SIMILATORY AND INHIBITORY ACTIONS ON THE RELEASE OF PITUITARY
CC GONADOTROPINS. THE EXACT FUNCTION OF THESE PROTEINS IS NOT KNOWN.
CC -1- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE II-LIKE DOMAINS.
CC -----
CC PIR: A29156; A29156.
CC HSSP: P02784; IPDC.
CC InterPro: IPR000562; FN_Type_II.
CC Pfam: PF00040; fn2; 2.
CC PRINTS: PR00013; FNYPEIT.
CC ProDom: PD000995; FN_Type_II; 2.
CC SMART: SM00059; FN2; 2.
CC PROSITE: PS00023; FIBRONECTIN_2; 2.
CC Senen; Plasma; Glycoprotein; Repeat.
CC FT DOMAIN 28 66 FIBRONECTIN TYPE-II, A.
CC FT DISULFID 29 53 FIBRONECTIN TYPE-II, B.
CC FT DISULFID 43 66 BY SIMILARITY.
CC FT DISULFID 74 100 BY SIMILARITY.
CC FT DISULFID 88 115 BY SIMILARITY.
CC SEQUENCE 115 AA; 13410 MW; 34E0FA871A9DE6 CRC64;
Query Match
Best Local Similarity 100.0%; Score 18; DB 1; Length 115;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLDA 4
Db 54 SLDA 57
RESULT 18
Y998_RHME ID Y998_RHME STANDARD: PRT: 116 AA.
AC Q52966;
DE 30-MAY-2000 (Rel. 39, Created)
DE 30-MAY-2000 (Rel. 39, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein R00998.

```

```

GN R00998 OR SMC00057.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
ON NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RCR2011 / SU47;
RX MEDLINE=95276304; PubMed=7756693;
RA Keller M., Koxlau A., Weng W.M., Schmidt M., Quandt J., Niehaus K.,
RT Jording D., Arnold W., Puhler A.;
RT "Molecular analysis of the Rhizobium meliloti mucr gene regulating the
RT biosynthesis of the exopolysaccharides succinoglycan and
RT galactoglucan.";
RL Mol. Plant Microbe Interact. 8:267-277(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetlelle D., Puhler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galbert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC -1- SIMILARITY: BELONGS TO THE UPPO091 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L37353; AAA74238.1; -.
DR EMBL; AL591785; CAC45570.1; -.
DR InterPro; IPR005133; Phag_mnhg_yufb.
DR Pfam; PF03334; Phag_mnhg_yufb.
DR TIGRfams; TIGR01300; CPA3_mnhg_phag; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 116 AA; 12373 MW; 16F4C117B81C4D10 CRC64;

Query Match 100.0%; Score 18; DB 1; Length 116;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4
Db 60 SLDA 63

RESULT 19
YE06_HAEIN STANDARD; PRT; 118 AA.
ID YE06_HAEIN
AC P4181;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein H11406.
GN H11406.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
ON NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,

```

```

RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Usterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Georghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U32820; AAC3056.1; -.
DR TIGR; H11406; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 118 AA; 13325 MW; C1CAF336CB3452 CRC64;

Query Match 100.0%; Score 18; DB 1; Length 118;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4
Db 101 SLDA 104

RESULT 20
CHH4_PENMO STANDARD; PRT; 120 AA.
ID CHH4_PENMO
AC 097386;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Crustacean hyperglycemic hormone 4 precursor (CHH) (Pm-SGP-IV).
GN CHH4.
OS Penaeus monodon (Penaeid shrimp).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata;
OC Penaeoidea; Penaeidae; Penaeus.
ON NCBI_TaxID=6687;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eyestalk;
RX PubMed=10804243;
RA Davey M.L., Hall M.R., Willis R.H., Oliver R.W.A., Thurn M.J.,
RA Wilson K.J.;
RT "Five crustacean hyperglycemic family hormones of Penaeus monodon:
RT complementary DNA sequence and identification in single sinus glands
RT by electrospray ionization-Fourier transform mass spectrometry.";
RL Mar. Biotechnol. 2:80-91(2000).
DECAPODS WHICH CONTROLS THE BLOOD SUGAR LEVEL. HAS A SECRETAGOGUE
ACTION OVER THE AMYLASE RELEASED FROM THE MIDGUT GLAND. MAY ACT AS
A STRESS HORMONE AND MAY BE INVOLVED IN THE CONTROL OF MOLTING AND
REPRODUCTION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE ARTHROPOD CHH/MH/SH/VIH FAMILY OF
CC HORMONES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL: AF104389; AAC84145.1; -.
DR InterPro: IPR001166; CHH_MiH_GH.
DR Pfam: PF01147; Crust_neurohorm. 1.
DR PRINTS: PR00550; HYPGLYCEMIC.
DR PROSITE: PS01250; CHH_MiH_GH; 1.
KW Neuropeptide; Hormone; Glucose metabolism; Amidation; Signal;
KW Cleavage on pair of basic residues; Multigene family.
FT SIGNAL 1 26
FT PEPTIDE 27 44 CHH PRECURSOR RELATED PEPTIDE (CRPP).
FT PEPTIDE 47 118 CRUSTACEAN HYPGLYCEMIC HORMONE 4.
FT DISULFID 53 89 BY SIMILARITY.
FT DISULFID 69 85 BY SIMILARITY.
FT DISULFID 72 98 BY SIMILARITY.
FT MOD_RES 118 118 AMIDATION (G-119 PROVIDE AMIDE GROUP) (BY
FT SIMILARITY).
SQ SEQUENCE 120 AA; 12987 MW; D04F39BDF7C4C1B9 CRC64;

Query Match
Best Local Similarity 100.0%; Score 18; DB 1; Length 120;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SIDA 4
   ||||
Db 28 SIDA 31

RESULT 21
RL22_HELPJ STANDARD; PRT; 122 AA.
ID RL22_HELPJ
AC O9ZJR8;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L22.
GN RPLV OR HPI1234.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moll D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tumano P.J., Carno A., Orla-Nielsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.
RL Nature 397:176-180(1999).
CC -1- FUNCTION: THIS PROTEIN BINDS SPECIFICALLY TO 23S RRNA: ITS BINDING
CC IS STIMULATED BY OTHER RIBOSOMAL PROTEINS, E.G., L4, L17, AND L20.
CC IT IS IMPORTANT DURING THE EARLY STAGES OF 50S RECONSTITUTION
CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE L22P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AE001547; AAD06800.1; -.
DR HSSP: P48286; IBXE.
DR InterPro: IPR001063; Ribosomal_L22.
DR Pfam: PF00237; Ribosomal_L22; 1.
DR ProDom: PD001032; Ribosomal_L22; 1.
DR TIGRfam: TIGR01044; rplV_bact; 1.
DR PROSITE: PS00464; RIBOSOMAL_L22; 1.
KW Ribosomal protein; rRNA-binding; Complete proteome.
SQ SEQUENCE 122 AA; 13048 MW; 63BE9285E51901523 CRC64;

```

```

Query Match
Best Local Similarity 100.0%; Score 18; DB 1; Length 122;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SIDA 4
   ||||
Db 58 SIDA 61

RESULT 22
RL22_HELPJ STANDARD; PRT; 122 AA.
ID RL22_HELPJ
AC P56047;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L22.
GN RPLV OR HPI1314.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kervavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klein H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson J., Dodson R., Khakh H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Meldrum J.M., Fujii C., Bowman C., Wathey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547(1997).
CC -1- FUNCTION: THIS PROTEIN BINDS SPECIFICALLY TO 23S RRNA: ITS BINDING
CC IS STIMULATED BY OTHER RIBOSOMAL PROTEINS, E.G., L4, L17, AND L20.
CC IT IS IMPORTANT DURING THE EARLY STAGES OF 50S RECONSTITUTION
CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE L22P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AE000633; AAD08353.1; -.
DR HSSP: P48286; IBXE.
DR TIGR: HPI1314;
DR InterPro: IPR001063; Ribosomal_L22.
DR Pfam: PF00237; Ribosomal_L22; 1.
DR ProDom: PD001032; Ribosomal_L22; 1.
DR TIGRfam: TIGR01044; rplV_bact; 1.
DR PROSITE: PS00464; RIBOSOMAL_L22; 1.
KW Ribosomal protein; rRNA-binding; Complete proteome.
SQ SEQUENCE 122 AA; 13076 MW; C3F02854DFB7895 CRC64;

Query Match
Best Local Similarity 100.0%; Score 18; DB 1; Length 122;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SIDA 4
   ||||
Db 58 SIDA 61

```



RESULT 23  
RS12\_XYLFA STANDARD: PRT: 124 AA.

AC Q9PAB8; 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE 30S ribosomal protein S12.  
GN RPSL\_OR XF2631.  
OS Xylella fastidiosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
OC Xylella.  
OX NCBI\_Taxid=2371;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=945C;  
RX MEDLINE=20365717; PubMed=10910347;  
RA Simpson A.J.G., Reinach F.C., Artuda P., Abreu F.A., Acencio M.,  
Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,  
Barrios M.H., Bonaccorsi E.D., Bordin S., Boye J.M., Briones M.R.S.,  
Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,  
Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,  
Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,  
Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,  
Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,  
Gantier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
Ho P.L., Hohnes J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,  
Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
Lemos E.G.M., Lemes M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,  
Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,  
Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
Nanni A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
de Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,  
Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,  
de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,  
Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,  
Zago M.A., Zaitz M., Meidanis J., Setubal J.C.;  
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";  
RL Nature 406:151-159(2000).  
CC -1- FUNCTION: PROTEIN S12 IS INVOLVED IN THE TRANSLATION INITIATION  
STEP (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE S12P FAMILY OF RIBOSOMAL PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: AE004069; AAF85428.1; ALT\_INIT.  
CC InterPro: IPR000230; Ribosomal\_S12.  
CC Pfam: PF00164; Ribosomal\_S12; 1.  
CC PRINTS: PR01034; Ribosomal\_S12.  
CC PRODOM: PD000576; Ribosomal\_S12; 1.  
CC TIGRfam: TIGR00981; rpsl\_bact; 1.  
CC PROSITE: PS00055; RIBOSOMAL\_S12; 1.  
CC RIBOSOMAL protein; Complete proteome.  
CC SQUENCE 124 AA; 13695 MW; 50211BD3A42C9FD1 CR64;

Query Match 100.0%; Score 18; DB 1; Length 124;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
||||

Db 101 SLDA 104

RESULT 24  
THN1\_WHEAT STANDARD: PRT: 126 AA.

AC P01544;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Alpha-1-purothionin precursor (purothionin A-II) (Fragment).  
GN TH1.1 OR PUR-D1.  
OS Triticum aestivum (Wheat).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
OC Triticeae; Triticum.  
OX NCBI\_Taxid=4565;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Endosperm;  
RX MEDLINE=95125120; PubMed=7824649;  
RA Castagnaro A., Marana C., Carbonero P., Garcia-Olmedo F.;  
RT "cDNA cloning and nucleotide sequences of alpha 1 and alpha 2  
RT thionins from hexaploid wheat endosperm.";  
RL Plant Physiol. 106:1221-1222(1994).  
RN [2]  
RP SEQUENCE OF 17-61.  
RC STRAIN=cv. Manitoba 3;  
RX MEDLINE=78026451; PubMed=914810;  
RA Ohtani S., Okada T., Yoshizumi H., Kagamiyama H.;  
RT "Complete primary structures of two subunits of purothionin A, a  
RT lethal protein for brewer's yeast from wheat flour.";  
RL J. Biochem. 82:753-767(1977).  
RN [3]  
RP SEQUENCE OF 17-61.  
RA Ohtani S., Okada T., Kagamiyama H., Yoshizumi H.;  
RT "The amino acid sequence of purothionin A, a lethal toxic protein  
RT to brewer's yeast from wheat.";  
RL Agric. Biol. Chem. 39:2269-2270(1975).  
RN [4]  
RP SEQUENCE OF 17-61.  
RC STRAIN=cv. Manitou;  
RA Jones B.L., Mak A.S.;  
RT "Amino acid sequences of the two alpha-purothionins of hexaploid  
RT wheat.";  
RL Cereal Chem. 54:511-523(1977).  
RN [5]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).  
RX MEDLINE=91045879; PubMed=2235992;  
RA Teeter M.M., Ma X.-Q., Rao U., Whitlow M.;  
RT "Crystal structure of a protein-toxin alpha 1-purothionin at 2.5A and  
RT a comparison with predicted models.";  
RL Proteins 8:118-132(1990).  
CC -1- FUNCTION: THIONINS ARE SMALL PLANT PROTEINS WHICH ARE TOXIC  
CC TO ANIMAL CELLS. THEY SEEM TO EXERT THEIR TOXIC EFFECT AT THE  
CC LEVEL OF THE CELL MEMBRANE. THE PRECISE FUNCTION, IN PLANTS,  
CC OF THESE PROTEINS IS NOT KNOWN.  
CC -----  
CC -1- SIMILARITY: BELONGS TO THE PLANT THIONIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X70666; CAA50004.1; -.  
CC PIR: A01807; VSWTA2.  
CC PDB: 2PDH; 03-APR-96.  
CC InterPro: IPR001010; Thionin.  
CC Pfam: PF00321; plant\_thionins; 1.  
CC PROSITE: PS00271; THIONIN; 1.

KW Thionin; Plant toxin; Signal; 3D-structure.  
 FT NON\_TER 1 1  
 FT SIGNAL <1 16  
 FT CHAIN 17 61 ALPHA-1-PUROTHIONIN.  
 FT CHAIN 62 126 ACIDIC PROTEIN.  
 FT DISULFID 19 55  
 FT DISULFID 20 47  
 FT DISULFID 28 45  
 FT DISULFID 32 41  
 SQ SEQUENCE 126 AA: 13525 MW: FF7310D921C4EE30 CRC64;

Query Match 100.0%; Score 18; DB 1; Length 126;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SIDA 4  
 DB 122 SIDA 125

RESULT 25  
 UL33\_HSV11 STANDARD; PRT; 130 AA.  
 ID UL33\_HSV11  
 AC P10217;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 01-DEC-1992 (Rel. 24, Last annotation update)  
 DE Protein UL33.  
 GN UL33.  
 OS Herpes simplex virus (type 1 / strain 17).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Alphaherpesvirinae; Simplexvirus.  
 OX NCBI\_TaxID=10299;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88274327; PubMed=2839594;  
 RA McGeoch D.J., Dalrymple M.A., Davison A.J., Dolan A., Frame M.C.,  
 RA McNab D., Perry L.J., Scott J.E., Taylor P.;  
 RT "The complete DNA sequence of the long unique region in the genome of  
 RT herpes simplex virus type 1.";  
 RL J. Gen. Virol. 69:1531-1574(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91082431; PubMed=1845831;  
 RA Al-Kobaisi M.F., Rixon F.J., McDougall I., Preston V.G.;  
 RT "The herpes simplex virus UL33 gene product is required for the  
 RT assembly of full capsids.";  
 RL Virology 180:380-388(1991).  
 CC -1- FUNCTION: STRUCTURAL; INVOLVED IN PACKAGING DNA.  
 CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL33,  
 CC EBV-1 27, AND VZV 25.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: D10879; BAA01679.1; -;  
 DR EMBL: X14112; CAA32308.1; -;  
 DR EMBL: M62932; AAA5829.1; -;  
 DR PIR: F30085; WWBEB3;  
 DR InterPro: IPR005208; Herpes\_UL33.  
 DR Pfam: PF03581; Herpes\_UL33; 1.  
 FT MUTAGEN 17 I->D: TEMPERATURE SENSITIVE.  
 SQ SEQUENCE 130 AA: 14437 MW: D04CFDA7A3C585D3 CRC64;

Query Match 100.0%; Score 18; DB 1; Length 130;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SIDA 4  
 DB 29 SIDA 32

RESULT 26  
 Y108\_MYCTU STANDARD; PRT; 133 AA.  
 ID Y108\_MYCTU  
 AC Q50717;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical protein RV3408.  
 DE RV3408 OR MT3516 OR MYCY78.20C.  
 GN Mycobacterium tuberculosis.  
 OS Mycobacteria; Actinobacteria (class); Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37Rv;  
 RX MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holtroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 RA Sutton J.E., Taylor K., Whitehead S., Barrett B.G.,  
 RA "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence.";  
 RL Nature 393:357-544(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CDC 1551 / Oshkosh;  
 RA Fleishmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
 RA Bishai W.;  
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
 RT laboratory strains.";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: Z77165; CAB01015.1; ALF\_INIT.  
 DR EMBL: AE007157; AAK47854.1; -;  
 DR TIGR: MT3516; -;  
 DR TubercuList: RV3408; -;  
 KW Hypothetical protein; Complete proteome.  
 FT CONFLICT 43 S->L (IN REF. 2).  
 SQ SEQUENCE 133 AA: 14294 MW: 4F3FE25FC816CF6B CRC64;

Query Match 100.0%; Score 18; DB 1; Length 133;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SIDA 4  
 DB 89 SIDA 92

RESULT 27  
 SFPL\_BOVIN STANDARD; PRT; 134 AA.  
 ID SFPL\_BOVIN

AC P02784;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Seminal plasma protein PDC-109 precursor (Seminal vesicle secretory  
 protein 109) (SVSP109) (BSP-A1 and BSP-A2).  
 OS Bos taurus (Bovine).  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;  
 CC Bovidae; Bovinae; Bos.  
 OC NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89152746; PubMed=3229283;  
 RA Kemme M., Scheit K.H.;  
 RT "Cloning and sequence analysis of a cDNA from seminal vesicle tissue  
 encoding the precursor of the major protein of bull semen.";  
 RL DNA 7:595-599(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RX MEDLINE=92031704; PubMed=1932121;  
 RA Breuer C., Scheit K.H.;  
 RT "Characterization of the gene for the bovine seminal vesicle  
 secretory protein SVSP109.";  
 RL Biochim. Biophys. Acta 1090:259-260(1991).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Breuer C.C., Kleine Kuhlmann J.J., Hanes J.J., Scheit K.K.;  
 RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE OF 26-134, AND DISULFIDE BONDS.  
 RX MEDLINE=83256590; PubMed=6870895;  
 RA Esch F.S., Ling N.C., Boehlen P., Ying S.Y., Gullermin R.;  
 RT "Primary structure of PDC-109, a major protein constituent of bovine  
 seminal plasma.";  
 RL Biochem. Biophys. Res. Commun. 113:861-867(1983).  
 RN [5]  
 RP CARBOHYDRATE-LINKAGE SITE THR-36.  
 RX MEDLINE=94350099; PubMed=8070564;  
 RA Calvete J.J., Ralda M., Sanz L., Wempe F., Scheit K.H., Romero A.,  
 RT Toepfer-Petersen E.;  
 RT "Localization and structural characterization of an oligosaccharide  
 O-linked to bovine PDC-109. Quantitation of the glycoprotein in  
 seminal plasma and on the surface of ejaculated and capacitated  
 spermatozoa.";  
 RL FEBS Lett. 350:203-206(1994).  
 RN [6]  
 RP STRUCTURE OF CARBOHYDRATE ON THR-36.  
 RX MEDLINE=96234019; PubMed=8654577;  
 RA Gerwig G.L., Calvete J.J., Toepfer-Petersen E., Vliegenhart J.F.G.;  
 RT "The structure of the O-linked carbohydrate chain of bovine seminal  
 plasma protein PDC-109 revised by H-NMR spectroscopy A correction.";  
 RL FEBS Lett. 387:99-100(1996).  
 RN [7]  
 RP STRUCTURE BY NMR OF 95-134.  
 RX MEDLINE=91129241; PubMed=1993183;  
 RA Constantine K.L., Ramesh V., Banyal L., Trexler M., Patthy L.,  
 RL Linas M.;  
 RT "Sequence-specific 1H NMR assignments and structural characterization  
 of bovine seminal fluid protein PDC-109 domain b.";  
 RL Biochemistry 30:1663-1672(1991).  
 RN [8]  
 RP STRUCTURE BY NMR OF 95-134.  
 RX MEDLINE=92114067; PubMed=1731074;  
 RA Constantine K.L., Madrid M., Banyal L., Trexler M., Patthy L.,  
 RL Linas M.;  
 RT "Refined solution structure and ligand-binding properties of PDC-109  
 domain b. A collagen-binding type II domain.";  
 RL J. Mol. Biol. 223:281-298(1992).  
 CC -I- FUNCTION: COULD ENHANCE THE FERTILIZING CAPACITY OF BULL  
 CC SPERMATOZOA UPON INTERACTION WITH HEPARIN-LIKE GLYCOSAMINOGLYCANS  
 CC PRESENT IN THE FEMALE GENITAL TRACT. EXHIBITS BOTH SIMULATORY AND

CC INHIBITORY ACTIONS ON THE RELEASE OF PITUITARY GONADOTROPINS.  
 CC -I- SUBUNIT: HOMODIMER.  
 CC -I- TISSUE SPECIFICITY: MAJOR COMPONENT OF SEMINAL PLASMA.  
 CC -I- PTM: O-GLYCANS CONSISTS OF GAL-GALNAc DISACCHARIDE WHICH IS  
 CC MODIFIED WITH A SULFIC ACID RESIDUE (MACRO-AND/OR  
 CC MICROHETEROGENEITY ACCOUNT FOR DIFFERENCES BETWEEN BSP-A1 AND BSP-  
 CC A2).  
 CC -I- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE II-LIKE DOMAINS.  
 CC -I- SIMILARITY: STRONG TO BSP-A3.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M22244; AAA30766.1; -;  
 DR EMBL; X60495; CAA43021.1; -;  
 DR EMBL; X60496; CAA43021.1; JOINED.  
 DR EMBL; X60497; CAA43021.1; JOINED.  
 DR EMBL; X60498; CAA43021.1; JOINED.  
 DR EMBL; X60960; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; Z33621; CAA83915.1; -;  
 DR PIR; A03256; WTBO.  
 DR PIR; A31852; A31852.  
 DR PIR; S18404; S18404.  
 DR PIR; S45010; S45010.  
 DR PDB; 1PDC; 31-JAN-94.  
 DR GlycoStatedB; P02784; -;  
 DR InterPro; IPR000562; FN\_Type\_II.  
 DR Pfam; PF00040; fn2; 2.  
 DR PRINTS; PR00013; FNTYPEII.  
 DR ProDom; PD000995; FN\_Type\_II; 2.  
 DR SMART; SM00059; FN2; 2.  
 DR PROSITE; PS00023; FIBRONECTIN\_2; 2.  
 KW Semen; Plasma; Glycoprotein; Repeat; signal; 3D-structure.  
 FT SIGNAL 1 25  
 FT CHAIN 26 134 SEMINAL PLASMA PROTEIN PDC-109.  
 FT DOMAIN 49 86 FIBRONECTIN TYPE-II, A.  
 FT 94 134 FIBRONECTIN TYPE-II, B.  
 FT CARBOHYD 36 36 O-LINKED (GALNAc. . .).  
 FT FTID-CAR\_000071.  
 FT DISULFID 49 73  
 FT DISULFID 63 86  
 FT DISULFID 94 119  
 FT DISULFID 108 134  
 FT STRAND 98 99  
 FT HELIX 100 102  
 FT STRAND 105 105  
 FT STRAND 109 109  
 FT TURN 111 112  
 FT STRAND 118 120  
 FT STRAND 124 124  
 FT TURN 126 127  
 FT TURN 131 133  
 FT SEQUENCE 134 AA; 15480 MW; AD5FE86969B32224 CRC64;  
 Query Match 100.0%; Score 18; DB 1; Length 134;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 SLDA 4  
 DB 74 SLDA 77  
 RESULT 28  
 ID YHIL\_MYCLE STANDARD; PRT; 134 AA.  
 AC P49774; Q9R734;  
 DT 01-OCT-1996 (Rel. 34, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE Hypothetical HIT-like protein ML2237.  
 GN ML2237 OR MLCB5.04C OR U296A.  
 OS Mycobacterium leprae.  
 CC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;  
 CC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 NX NCBI\_TaxID=1169;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Smith D.R., Robison K.;  
 RL Submitted (SEP-1994) to the EMBL/Genbank/DBD databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-TN;  
 RX MEDLINE=21128732; PubMed=11234002;  
 RA Cole S.T., Eglmeier K., Parkhill J., James K.D., Thomson N.R.,  
 RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,  
 RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,  
 RA Holroyd S., Hornsby T., Jagsels K., Lacroix C., Maclean J., Moule S.,  
 RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,  
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,  
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,  
 RA Barrell B.G.;  
 RT "Massive gene decay in the leprosy bacillus."  
 RL Nature 409:1007-1011(2001).  
 CC -1- SIMILARITY: BELONGS TO THE HIT FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U5187; AAA63136.1; ALT\_INT.  
 DR EMBL: Z95151; CAB08415.1; -;  
 DR EMBL: AF583924; CAC31193.1; -;  
 DR HSP: P80912; 4RHN.  
 DR Leproma; ML2237; -;  
 DR InterPro: IPR001310; HIT.  
 DR Pfam: PF01230; HIT; 1.  
 DR PRINTS: PR00332; HISTRIAD.  
 DR PROSITE: PS00892; HIT; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 134 AA; 14712 MW; C1853F3ADF7239A7 CRC64;  
 Query Match 100.0%; Score 18; DB 1; Length 134;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 SIDA 4  
 DB 118 SIDA 121  
 RESULT 29  
 YNP0\_YEAST  
 ID YNP0\_YEAST STANDARD; PRT; 135 AA.  
 AC P53902;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Hypothetical 14.8 kDa protein in RPC8-MFA2 intergenic region.  
 GN YNL150W OR N1773.  
 OS Saccharomyces cerevisiae (Baker's Yeast).  
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycos.  
 NX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-S288C / FY1679;  
 RX MEDLINE=96287653; PubMed=8686380;  
 RA Nasr F., Becam A.-M., Herbert C.J.;  
 RT "The sequence of 36.8 kb from the left arm of chromosome XIV reveals  
 RT 24 complete open reading frames: 18 correspond to new genes, one of  
 RT which encodes a protein similar to the human myotonic dystrophy  
 RT kinase."  
 RL Yeast 12:169-175(1996).  
 RN [2]  
 RP SEQUENCE OF 104-135 FROM N.A.  
 RC STRAIN-S288C;  
 RX MEDLINE=96109932; PubMed=8619318;  
 RA Mallet L., Bussereau F., Jacquet M.;  
 RT "A 43.5 kb segment of yeast chromosome XIV, which contains MFA2,  
 RT MEP2, CAP/SRV2, NAM9, FKBI/FPRI/RPI1, MOM22 and CPT1, predicts an  
 RL yeast 11:1195-1209(1995).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X92517; CAA63289.1; -;  
 DR EMBL: Z71424; CAA96032.1; -;  
 DR EMBL: Z71426; CAA96036.1; -;  
 DR SGD: S0005094; YNL150W.  
 KW Hypothetical protein.  
 FT DOMAIN 41  
 SQ SEQUENCE 135 AA; 14844 MW; 22EC1EG99F3DDA5 CRC64;  
 Query Match 100.0%; Score 18; DB 1; Length 135;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 SIDA 4  
 DB 126 SIDA 129  
 RESULT 30  
 THN2\_WHEAT  
 ID THN2\_WHEAT STANDARD; PRT; 136 AA.  
 AC P32032;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE Alpha-2-purothionin precursor.  
 GN THN1.2 OR PUR-BL.  
 OS Triticum aestivum (Wheat).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;  
 CC Triticeae; Triticum.  
 NX NCBI\_TaxID=4565;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Endosperm;  
 RX MEDLINE=95125120; PubMed=7824649;  
 RA Castagnaro A., Marana C., Carbonero P., Garcia-Olmedo F.;  
 RT "cDNA cloning and nucleotide sequences of alpha 1 and alpha 2  
 RT thionins from hexaploid wheat endosperm."  
 RL Plant Physiol. 106:1221-1222(1994).  
 CC -----  
 CC -1- FUNCTION: THIONINS ARE SMALL PLANT PROTEINS WHICH ARE TOXIC  
 CC TO ANIMAL CELLS. THEY SEEM TO EXERT THEIR TOXIC EFFECT AT THE  
 CC LEVEL OF THE CELL MEMBRANE. THE PRECISE FUNCTION, IN PLANTS,  
 CC OF THESE PROTEINS IS NOT KNOWN.  
 CC -1- SIMILARITY: BELONGS TO THE PLANT THIONIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X70665; CA50003.1; -  
 DR PIR: S31695; S31695.  
 DR HSSP: P01543; IBHP.  
 DR InterPro: IPR001010; Thionin.  
 DR Pfam: PF00321; Plant\_Thionins; 1.  
 DR PROSITE: PS00271; THIONIN; 1.  
 KM Thionin; Plant toxin; Signal.  
 FT SIGNAL 1 27  
 FT CHAIN 28 72 ALPHA-2-PUROTHIONIN.  
 FT DISULFID 73 136 ACIDIC PROTEIN.  
 FT DISULFID 30 66 BY SIMILARITY.  
 FT DISULFID 31 58 BY SIMILARITY.  
 FT DISULFID 39 56 BY SIMILARITY.  
 FT DISULFID 43 52 BY SIMILARITY.  
 SQ SEQUENCE 136 AA; 14558 MM; B4019F014E226B9F CRC64;  
 Query Match 100.0%; Score 18; DB 1; Length 136;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 SLDA 4  
 Db 133 SLDA 136  
 RESULT 31  
 THNB\_WHEAT  
 ID THNB\_WHEAT STANDARD; PRT; 136 AA.  
 AC P01543;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Purothionin A-I precursor (Beta-purothionin).  
 GN TH11.3.  
 OS Triticum aestivum (Wheat).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae;  
 OC Triticaceae; Triticum.  
 OX NCBI\_TaxID=4565;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Rosella;  
 RA Hughes P.A., Llewellyn D.L., Whitecross M.;  
 RL Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 28-72.  
 RC STRAIN=cv. Manitoba 3;  
 RX MEDLINE=78026451; PubMed=914810;  
 RA Ohtani S., Okada T., Yoshizumi H., Kagamiyama H.;  
 RT "Complete primary structures of two subunits of purothionin A, a  
 RT lethal protein for brewer's yeast from wheat flour.";  
 RL J. Biochem. 82:753-767(1977).  
 RN [3]  
 RP SEQUENCE OF 28-72.  
 RA Ohtani S., Okada T., Kagamiyama H., Yoshizumi H.;  
 RT "The amino acid sequence of purothionin A, a lethal toxic protein  
 RT to brewer's yeast from wheat".  
 RL Agric. Biol. Chem. 39:2269-2270(1975).  
 RN [4]  
 RP SEQUENCE OF 28-72.  
 RX MEDLINE=77046666; PubMed=990986;  
 RA Mak A.S., Jones B.L.;  
 RT "The amino acid sequence of wheat beta-purothionin.";  
 RL Can. J. Biochem. 54:835-842(1976).  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).  
 RA Stec B., Rao U., Teeter M.M.;

RT "Refinement of purothionins reveals solute particles important for  
 RT lattice formation and toxicity. Part 2: structure of beta-purothionin  
 RT at 1.7-A resolution.";  
 RL Acta Crystallogr. D 51:914-924(1995).  
 CC -1- FUNCTION: THIONINS ARE SMALL PLANT PROTEINS WHICH ARE TOXIC  
 CC TO ANIMAL CELLS. THEY SEEM TO EXERT THEIR TOXIC EFFECT AT THE  
 CC LEVEL OF THE CELL MEMBRANE. THE PRECISE FUNCTION, IN PLANTS,  
 CC OF THESE PROTEINS IS NOT KNOWN.  
 CC -1- SIMILARITY: BELONGS TO THE PLANT THIONIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF004018; AAB71137.1; -  
 DR PIR: A01806; VSWTAL.  
 DR PDB: IBHP; 15-MAR-96.  
 DR InterPro: IPR001219; Neurotoxin.  
 DR InterPro: IPR001010; Thionin.  
 DR Pfam: PF00321; Plant\_Thionins; 1.  
 DR PRINTS: PR00284; TOXIN.  
 DR PROSITE: PS00271; THIONIN; 1.  
 KM Thionin; Plant toxin; Signal; 3D-structure.  
 FT SIGNAL 1 27  
 FT CHAIN 28 72  
 FT CHAIN 73 136 PUROTHIONIN A-I.  
 FT DISULFID 30 66 ACIDIC PROTEIN.  
 FT DISULFID 31 58  
 FT DISULFID 39 56  
 FT DISULFID 43 52  
 SQ SEQUENCE 136 AA; 14625 MM; A855C81519EDA24 CRC64;  
 Query Match 100.0%; Score 18; DB 1; Length 136;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 SLDA 4  
 Db 133 SLDA 136  
 RESULT 32  
 NNT1\_CAEEL  
 ID NNT1\_CAEEL STANDARD; PRT; 137 AA.  
 AC Q9U757;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE NTF2-related export protein.  
 GN NNT1 OR Y71F9AM.5.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20036817; PubMed=10567585;  
 RA Black B.E., Levesque L., Holaska J.M., Wood T.C., Paschal B.M.;  
 RT "Identification of an NTF2-related factor that binds Ran-GTP and  
 RT regulates nuclear protein export.";  
 RL Mol. Cell. Biol. 19:8616-8624(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Bradshaw-Cordum H., Scott K., Graves T.;  
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: Stimulator of protein export for NES-containing  
 CC proteins. Also plays a role in the nuclear export of. U1 snRNA,  
 CC tRNA, and mRNA (by similarity).

```

CC -1- SUBUNIT: Preferentially binds Ran-GTP (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: CONTAINS 1 NTF2 DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: AF156960; AAD54945.1; -
DR EMBL: AC024852; AAK66028.1; -
DR HSSP: P13662; IOUN.
DR WormPep: Y71F9AM.5; CE26780.
DR InterPro: IPR002075; NTF2.
DR Pfam: PF02136; NTF2; 1.
DR PROSITE: PS50177; NTF2_DOMAIN; 1.
KW Transport; Protein transport; mRNA transport; Nuclear protein.
FT DOMAIN
SQ SEQUENCE 137 AA; 15479 MW; B99ABE3CA3479962 CRC64;

Query Match
Best Local Similarity 100.0%; Score 18; DB 1; Length 137;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4
   ||||
Db 77 SLDA 80

RESULT 33
NUSL_YERPE STANDARD: PRT; 138 AA.
AC Q8ZC42;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE N utilization substance protein B homolog (NusB protein).
DE NUSB OR YPO3181.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
OX NCBI_TaxID=632;
[1]
RN SEQUENCE FROM N.A.
RP SRRAIN-CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Farrago A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Littlewell T., Hamlin N., Holroyd S., Jagers K., Kariyev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford A.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrett B.G.;
RA Genome sequence of Yersinia pestis, the causative agent of plague.
RT Nature 413:523-527(2001).
RL -1- FUNCTION: Involved in the transcription termination process (By
similarity).
CC -1- SIMILARITY: BELONGS TO THE NUSB FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: A0414155; CAC92416.1; -
DR InterPro: IPR000139; NusB.
DR Pfam: PF01029; NusB; 1.

```

```

KW Transcription termination; Complete proteome.
SQ SEQUENCE 138 AA; 15506 MW; 71A9A0307E6473A9 CRC64;

Query Match
Best Local Similarity 100.0%; Score 18; DB 1; Length 138;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4
   ||||
Db 61 SLDA 64

RESULT 34
PSB2_ECOLI STANDARD: PRT; 144 AA.
ID PSB2_ECOLI
AC P10032;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE PSIB protein.
DE PSIB.
GN Escherichia coli.
OS Plasmid R6-5.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=89083490; PubMed=3205720;
RA Duteilh M., Baeckman A., Celerier J., Bagdasarian M., Sommer S.,
RA Ballone A., Devoret R., Bagdasarian M.;
RT Identification of psib genes of plasmids F and R6-5. Molecular basis
RT for psib enhanced expression in plasmid R6-5.
RL Nucleic Acids Res. 16:10669-10679(1988).
CC -1- FUNCTION: INHIBITION OF THE INDUCTION OF THE SOS PATHWAY.
CC -1- SIMILARITY: STRONG, TO PSIB PROTEIN OF PLASMID F.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: X12463; CAA31004.1; -
DR DR PIR: S01897; S01897.
DR KM Plasmid.
SQ SEQUENCE 144 AA; 15777 MW; 369A28FB96B0887 CRC64;

Query Match
Best Local Similarity 100.0%; Score 18; DB 1; Length 144;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4
   ||||
Db 122 SLDA 125

RESULT 35
HFPA_CAUCR STANDARD: PRT; 147 AA.
ID HFPA_CAUCR
AC P27342; 068957;
DT 01-AUG-1992 (Rel. 23, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE HflaA attachment protein A (HfaA protein) (Hfa-A).
DE HFPA OR CC2628.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=155892;
[1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-CB2;
RX MEDLINE-92121105; PubMed-1732204;
RA Kurtz H.D. Jr., Smit J.;
RT "Analysis of a Caulobacter crescentus gene cluster involved in
RL attachment of the holdfast to the cell."
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 19089 / CB15;
RX MEDLINE-99138730; PubMed-9973336;
RA Usankarman R.S., Brun Y.V.;
RT "Cell cycle control of a holdfast attachment gene in Caulobacter
RL crescentus."
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 19089 / CB15;
RX MEDLINE-21173698; PubMed-11259647;
RA Nierman W.C., Feldblum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton M.S., Gwinn M.L., Haft D.H.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Ullrich T., Tran K., Wolf A., Yamachyan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001)
CC -!- FUNCTION: INVOLVED IN ATTACHMENT OF THE HOLDFAST TO THE CELL. THE
CC HOLDFAST IS A STRUCTURE THAT ALLOWS THE BACTERIA TO FIRMLY
CC ADHERE TO SURFACES.
CC -!- SIMILARITY: SOME, IN THE C-TERMINUS WITH E. COLI PARG AND
CC S. MARCESCENS SMGC C-TERMINI.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M69129; AAA23044.1; -
DR EMBL; AF058792; AAC14298.1; -
DR EMBL; AE005930; AAK24596.1; -
DR PIR; A42610; A42610.
DR TIGR; CC2628; -.
DR KW Complete proteome.
FT CONFLICT 23 S -> C (IN REF. 1).
FT CONFLICT 41 G -> R (IN REF. 2).
FT CONFLICT 78 S -> N (IN REF. 2).
FT CONFLICT 125 NNSOV -> IORGE (IN REF. 1).
SQ SEQUENCE 147 AA; 14699 MW; 09B10FE5DF436F CRC64;

Query Match 100.0%; Score 18; DB 1; Length 147;
Best Local Similarity 100.0%; Pred. No. 3,1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SIDA 4
Db 88 SIDA 91

RESULT 36
RS12_SULSO STANDARD; PRT; 147 AA.
AC P39573;
DT 01-FEB-1995 (Rel. 31, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 30S ribosomal protein S12P.
GN RPS12P OR RPS12 OR RPS12AB OR SSO0219.
OS Sulfolobus solfataricus.

```

```

OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 49255 / DSM 5833 / MT-4;
RX MEDLINE-94198299; PubMed-8148382;
RA Arcari P., Gallo M., Iannicelli G., Dello Russo A., Bocchini V.;
RT "The nucleotide sequence of the gene coding for the elongation factor
RL 1 alpha in Sulfolobus solfataricus. Homology of the product with
RN related proteins."
RL Biochim. Biophys. Acta 1217:333-337(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 35092 / DSM 1617 / P2;
RX MEDLINE-21332296; PubMed-11427726;
RA She Q., Singh R.K., Contaloni F., Zivanovic Y., Allard G.,
RA Amayez M.J., Chan-Welher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Erasmo G., Fletcher C., Gordon P.M.R.,
RA Helkamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngo H.P., Redder P., Schenk M.E., Thieriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2."
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
CC -!- FUNCTION: PROTEIN S12 IS INVOLVED IN THE TRANSLATION INITIATION
CC STEP.
CC -----
CC -!- SIMILARITY: BELONGS TO THE S12P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X76767; CAA54160.1; -
DR EMBL; AE006659; AAK40562.1; ALT. INIT.
DR InterPro; IPR000230; Ribosomal_S12.
DR Pfam; PF00164; Ribosomal_S12; 1.
DR ProDom; PD000576; Ribosomal_S12; 1.
DR TIGRFAMs; TIGR00982; S23_S12_E_A; 1.
DR PROSITE; PS00055; RIBOSOMAL_S12; 1.
KW Ribosomal protein; Complete proteome.
FT VARIANT 2 V -> S (IN STRAIN MT-4).
FT VARIANT 7 P -> S (IN STRAIN MT-4).
FT VARIANT 12 A -> S (IN STRAIN MT-4).
FT VARIANT 42 D -> N (IN STRAIN MT-4).
FT CONFLICT 128 MV -> NG (IN REF. 1).
SQ SEQUENCE 147 AA; 16346 MW; 740BF81849BDE32 CRC64;

Query Match 100.0%; Score 18; DB 1; Length 147;
Best Local Similarity 100.0%; Pred. No. 3,1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SIDA 4
Db 133 SIDA 136

RESULT 37
RS12_SULTO STANDARD; PRT; 147 AA.
AC Q976A8;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S12P.
GN RPS12P OR ST0272.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.

```

```

OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RA MEDLINE=21456156; PubMed=11572479;
RA Kawarayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kuchō Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagi M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermophilic
RT Cremonarchaeon, Sulfolobus tokodaii strain 7."
RL DNA Res. 8:123-140(2001).
CC -1- FUNCTION: Protein S12 is involved in the translation initiation
CC step (by similarity).
CC -1- SIMILARITY: BELONGS TO THE S12P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AP000962; BAB5239.1; -
DR InterPro: IPR000230; Ribosomal_S12.
DR Pfam: PF00164; Ribosomal_S12; 1.
DR ProDom: PD000576; Ribosomal_S12; 1.
DR TIGRfams: TIGR00982; S23_S12_E.A; 1.
DR PROSITE: PS00055; RIBOSOMAL_S12; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 147 AA; 16336 MW; 5BC87B93A599A1FF CRC64;

Query Match
Best Local Similarity 100.0%; Score 18; DB 1; Length 147;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4
DB 133 SLDA 136

RESULT 38
TPCB_HOMAM STANDARD; PRT; 150 AA.
AC P29291;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE Troponin C, isoform 2B
OS Homarus americanus (American lobster).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;
OC Astacidae; Nephropoidea; Nephropidae; Homarus.
OX NCBI_TaxID=6706;
RN [1]
RP SEQUENCE.
RC TISSUE=Abdominal flexor muscle;
RX MEDLINE=92027802; PubMed=1929438;
RA Garone L., Theibert J.L., Miegel A., Maeda Y., Murphy C.,
RA Collins J.H.;
RT "Lobster troponin C: amino acid sequences of three isoforms."
RL Arch. Biochem. Biophys. 291:89-91(1991).
CC -1- FUNCTION: TROPONIN IS THE CENTRAL REGULATORY PROTEIN OF STRIATED
CC MUSCLE CONTRACTION. TN CONSISTS OF THREE COMPONENTS: TN-I WHICH
CC IS THE INHIBITOR OF ACTOMYOSIN ATPASE, TN-T WHICH CONTAINS THE
CC BINDING SITE FOR TROPOMYOSIN AND TN-C, THE BINDING OF CALCIUM TO
CC TN-C ABOLISHES THE INHIBITORY ACTION OF TN ON ACTIN FILAMENTS.
CC -1- MISCELLANEOUS: THERE ARE THREE DIFFERENT TROPONIN C IN LOBSTER.
CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.

```

```

DR PIR, S18435; S18435.
DR HSSP; P02588; IAVS.
DR InterPro: IPR002048; EF-hand.
DR Pfam: PF00036; efhand; 4.
DR ProDom: PD000012; EF-hand; 2.
DR SMART; SM00054; Eph; 4.
DR PROSITE; PS00018; EF_HAND; 2.
KW Muscle protein; Calcium-binding; Repeat; Acetylation.
FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
FT DOMAIN 23 34 ANCESTRAL CALCIUM SITE 1 (POTENTIAL).
FT CA_BIND 56 67 EF-HAND 2 (PROBABLE).
FT DOMAIN 96 107 ANCESTRAL CALCIUM SITE 3 (POTENTIAL).
FT CA_BIND 132 143 EF-HAND 4 (PROBABLE).
SQ SEQUENCE 150 AA; 16951 MW; 56A0FA305C7D5F42 CRC64;

Query Match
Best Local Similarity 100.0%; Score 18; DB 1; Length 150;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4
DB 3 SLDA 6

RESULT 39
MRAZ_PSEAE STANDARD; PRT; 151 AA.
AC Q9HYZ4;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protein mraz.
DE Mraz.
GN MRAZ OR PA4421.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltz L., Tolentino E., Westbrook-Wadman S., Yan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
CC -1- SIMILARITY: BELONGS TO THE MRAZ FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE004856; AAC07809.1; -
DR InterPro: IPR003444; UPE0040.
DR Pfam: PF02381; UPE0040; 2.
DR TIGRfams: TIGR00242; TIGR00242; 1.
KW Complete proteome.
SQ SEQUENCE 151 AA; 17111 MW; 2E653E29F113216B CRC64;

Query Match
Best Local Similarity 100.0%; Score 18; DB 1; Length 151;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4
DB 133 SLDA 136

```



Db 9 SLDA 12

RESULT 40

RS12.SULAC STANDARD: PRT; 151 AA.

ID RS12.SULAC

AC P11524;

DT 01-OCT-1989 (Rel. 12, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE 30S ribosomal protein S12P.

GN RPS12P OR RPS12.

OS Sulfolobus acidocaldarius.

OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;

CC Sulfolobus.

CC NCBI\_TaxID=2285;

OX NCBI\_TaxID=2285;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-ATCC 33909 / NCIB 11770 / DSM 639;

RX MEDLINE=89315197; PubMed=2501756;

RA Puehler G., Lottspeich F., Zillig W.;

RT "Organization and nucleotide sequence of the genes encoding the large subunits A, B and C of the DNA-dependent RNA polymerase of the archaeobacterium Sulfolobus acidocaldarius.";

RT Nucleic Acids Res. 17:4517-4534(1989).

RL

CC -1- FUNCTION: PROTEIN S12 IS INVOLVED IN THE TRANSLATION INITIATION STEP.

CC

CC -1- SIMILARITY: BELONGS TO THE S12P FAMILY OF RIBOSOMAL PROTEINS.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

CC EMBL: X14818; CAA32929.1; ALT\_INIT.

DR PIR: S04721; R3UC12.

DR InterPro: IPR00230; Ribosomal\_S12.

DR Pfam: PF00164; Ribosomal\_S12; 1.

DR ProDom: PD000576; Ribosomal\_S12; 1.

DR TIGRFAMs: TIGR00982; S23\_S12\_E\_A; 1.

DR PROSITE: PS00055; RIBOSOMAL\_S12; 1.

KW Ribosomal protein.

SQ SEQUENCE 151 AA; 16570 MW; 3867AE02D161491F CRC64;

Query Match 100.0%; Score 18; DB 1; Length 151;

Best Local Similarity 100.0%; Pred. No. 3.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4

Db 133 SLDA 136

RESULT 41

Y4JR.RHISN STANDARD: PRT; 151 AA.

ID Y4JR.RHISN

AC P55518;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Hypothetical 17.3 kDa protein Y4JR.

GN Y4JR.

OS Rhizobium sp. (strain NGR234).

OC Plasmid sym pNGR234e.

CC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

CC Rhizobiaceae; Rhizobium.

OX NCBI\_TaxID=394;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=97305956; PubMed=9163424;

RA Freiberg C.A., Fellay R., Balroch A., Broughton W.J., Rosenthal A., Perret X.;

RT "Molecular basis of symbiosis between Rhizobium and legumes.";

RL Nature 387:394-401(1997).

CC

CC -1- SIMILARITY: CONTAINS 1 RESPONSE REGULATORY DOMAIN.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

CC EMBL: AE00080; AAB91730.1; -.

DR InterPro: IPR001789; Response\_reg.

DR SMART: SM00448; REC; 1.

DR PROSITE: PS0110; RESPONSE\_REGULATORY; 1.

KW Hypothetical protein; Plasmid.

FT DOMAIN 1

SQ SEQUENCE 151 AA; 17262 MW; F8B42BCF7845E03F CRC64;

Query Match 100.0%; Score 18; DB 1; Length 151;

Best Local Similarity 100.0%; Pred. No. 3.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4

Db 112 SLDA 115

RESULT 42

COBE.PSEDE STANDARD: PRT; 154 AA.

ID COBE.PSEDE

AC P21635;

DT 01-MAY-1991 (Rel. 18, Created)

DT 01-MAY-1991 (Rel. 18, Last sequence update)

DT 01-JUL-1993 (Rel. 26, Last annotation update)

DE COBE protein.

GN COBE.

OS Pseudomonas denitrificans.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

CC Pseudomonas.

OX NCBI\_TaxID=43306;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SC510;

RX MEDLINE=91008975; PubMed=2211520;

RA Crouzet J., Cauchols L., Blanche F., Debussche L., Thibaut D., Rouyez M.-C., Rigault S., Mayaux J.-F., Cameron B.;

RT "Nucleotide sequence of a pseudomonas denitrificans 5.4-kilobase DNA fragment containing five cob genes and identification of structural genes encoding S-adenosyl-L-methionine: uroporphyrinogen III methyltransferase and cobyrinic acid a,c-diamide synthase.";

RT J. Bacteriol. 172:5968-5979(1990).

RL

CC -1- PATHWAY: COBALAMIN BIOSYNTHESIS; TRANSFORMATION OF PRECORIN-2 TO COBRINIC ACID.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

CC EMBL: M59236; AAA25772.1; -.

DR PIR: A36144; A36144.

DR InterPro: IPR002750; CbiG.

DR Pfam: PF01890; CbiG; 1.

KW Cobalamin biosynthesis; Porphyrin biosynthesis.

SQ SEQUENCE 154 AA; 15487 MW; 197D77DC4CE20BAE CRC64;

Query Match 100.0%; Score 18; DB 1; Length 154;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
 ||||  
 DB 55 SLDA 58

RESULT 43  
 VATL\_HELVIT STANDARD; PRT; 156 AA.

AC P55277; 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DE 15-JUN-2002 (Rel. 41, Last annotation update)  
 DR Vacuolar ATP synthase 16 kDa proteolipid subunit (EC 3.6.3.14).  
 GN VHA16  
 OS Heliothis virescens (Noctuid moth) (Owlet moth).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 OC Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;  
 OC Ditrysia; Noctuoidea; Noctuidae; Heliothinae; Heliothis.  
 OX NCBI\_TaxID=7102;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC TISUP-Midgut.  
 RX MEDLINE=93357792; PubMed=8353524;  
 RA Pietrantoni P.V., Gill S.S.;

"Sequence of a 17 kDa vacuolar H(+) -ATPase proteolipid subunit from  
 insect midgut and Malpighian tubules."  
 RT Insect Biochem. Mol. Biol. 23:675-680(1993).

RL -1 FUNCTION: PROTON-CONDUCTING PORE FORMING. SUBUNIT OF THE MEMBRANE  
 INTEGRAL V0 COMPLEX OF VACUOLAR ATPASE. V-ATPASE IS RESPONSIBLE  
 FOR ACIDIFYING A VARIETY OF INTRACELLULAR COMPARTMENTS IN  
 EUKARYOTIC CELLS.

CC -1 CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +  
 H(+) (out).

CC -1 SUBUNIT: V-ATPASE IS AN HETEROMULTIMERIC ENZYME COMPOSED OF A  
 PERIPHERAL CATALYTIC V1 COMPLEX (MAIN COMPONENTS: SUBUNITS A, B,  
 C, D, E, AND F) ATTACHED TO AN INTEGRAL MEMBRANE V0 PROTON PORE  
 COMPLEX (MAIN COMPONENT: THE PROTEOLIPID PROTEIN) WHICH IS PRESENT  
 AS A HEXAMER THAT FORMS THE PROTON-CONDUCTING PORE.

CC -1 SUBCELLULAR LOCATION: Integral membrane protein. Vacuolar.  
 CC -1 MISCELLANEOUS: THIS SUBUNIT BINDS DICYCLOHEXYLCARBODIIMIDE (DCCD)  
 WHICH INHIBITS THE ATPASE (BY SIMILARITY).

CC -1 SIMILARITY: BELONGS TO THE V-ATPASE PROTEOLIPID SUBUNIT FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: L16884; AAC37176.1; -  
 DR InterPro: IPR002379; ATPase\_Csub.  
 DR InterPro: IPR00245; Vac\_ATPsyn\_Csub.  
 DR Pfam: PF00137; ATP-synt\_C; 2.  
 DR PRINTS: PR00122; VACATPASE.

DR TIGRfam: TIGR01100; V\_ATP\_synt\_C; 1  
 KM Hydrolyase; Hydrogen ion transport; ATP synthesis; Transmembrane.  
 FT DOMAIN 1  
 FT TRANSSEM 8  
 FT TRANSSEM 30  
 FT TRANSSEM 31  
 FT TRANSSEM 52  
 FT TRANSSEM 53  
 FT TRANSSEM 73  
 FT TRANSSEM 74  
 FT TRANSSEM 92  
 FT TRANSSEM 93  
 FT TRANSSEM 114  
 FT TRANSSEM 126  
 FT TRANSSEM 127  
 FT TRANSSEM 156  
 FT BINDING 139  
 FT SEQUENCE 156 AA; 15934 MW; E8DB1065B896E4B5 CRC64;

Query Match 100.0%; Score 18; DB 1; Length 156;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
 ||||  
 DB 76 SLDA 79

RESULT 44  
 UL51\_HCMVA STANDARD; PRT; 157 AA.

AC P16792; 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DE 01-FEB-1991 (Rel. 17, Last annotation update)  
 DR Hypothetical protein UL51.  
 GN UL51.  
 OS Human cytomegalovirus (strain AD169).

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Betaherpesvirinae; Cytomegalovirus.  
 OX NCBI\_TaxID=10360;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC Chee M.S., Bankier A.T., Beck S., Bohm R., Brown C.M., Cerny R.,  
 Horsnell T., Hutchinson C.A., Ilt, Kouzarides T., Martignetti J.A.,  
 RA Pridell E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;

"Analysis of the protein-coding content of the sequence of human  
 cytomegalovirus strain AD169."  
 RT Curr. Top. Microbiol. Immunol. 154:125-169(1990).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: X17403; CA35410.1; -  
 DR PIR: S09814; S09814.  
 DR InterPro: IPR005208; Herpes\_UL33.  
 DR Pfam: PF03581; Herpes\_UL33; 1.

DR Hypothetical protein.  
 KM SEQUENCE 157 AA; 16968 MW; 5999035AC654517D CRC64;

Query Match 100.0%; Score 18; DB 1; Length 157;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
 ||||  
 DB 92 SLDA 95

RESULT 45  
 RISB\_MYCTU STANDARD; PRT; 160 AA.

AC P1685; 01-NOV-1997 (Rel. 35, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE 6,7-dimethyl-8-ribityllumazine synthase (EC 2.5.1.9) (DMRL synthase)  
 DE (lumazine synthase) (Riboflavin synthase beta chain).  
 GN RIBH OR RV1416 OR MT1459 OR MTCY21B4.34.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;  
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.

OX NCBI\_TaxID=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RX STRAIN=H37RV: PubMed=9634230;  
 RA MEDLINE=98295987; Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Cole S.T., Brosch R., Parkhill J., Garon L., Barry C.E. III, Tekala F.,  
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;  
 RA "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence."  
 RL Nature 393:537-544(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CDC 1551 / Oshkosh;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., Deboy R., Dodson R., Gryn M.L., Haft D., Hickey E.,  
 RA Kolony J.F., Nelson W.C., Umayam L.A., Ernolova M.D., Salzberg S.L.,  
 RA Delcher A., Utterback T., Weidman J., Kouri H., Gill J., Mikula A.,  
 RA Bishai W.;  
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
 RT laboratory strains."  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Riboflavin synthase is a bifunctional enzyme complex  
 CC catalyzing the formation of riboflavin from 5-amino-6-(1'-D)-  
 CC ribityl-amino-2,4(1H,3H)-pyrimidinone and L-3,4-dihydroxy-2-  
 CC butanone-4-phosphate via 6,7-dimethyl-8-lumazine. The beta subunit  
 CC catalyzes the condensation of 5-amino-6-(1'-D)-ribityl-amino-  
 CC 2,4(1H,3H)-pyrimidinone with L-3,4-dihydroxy-2-butanone-4-  
 CC phosphate yielding 6,7-dimethyl-8-lumazine (By similarity).  
 CC -1- CATALYTIC ACTIVITY: 2,6,7-dimethyl-8-(1'-D-ribityl)lumazine -  
 CC riboflavin + 4-(1'-D-ribitylamino)-5-amino-2,6-dihydroxypyrimidine.  
 CC -1- PATHWAY: Riboflavin biosynthesis: last step.  
 CC -1- SIMILARITY: BELONGS TO THE DMRL SYNTHASE FAMILY.  
 CC -----  
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Z80108; CAB02164.1; ALT.INIT.  
 DR EMBL: AE007016; AAK45724.1; -.  
 DR HSSP: O65529; IHOK.  
 DR TIGR: MT1459; -.  
 DR TubercuList: RV1416; -.  
 DR InterPro: IPR002180; DMRL\_synthase.  
 DR Pfam: PF00885; DMRL\_synthase; 1.  
 DR ProDom: PD003664; DMRL\_synthase; 1.  
 DR Trifam: TIGR00114; ribH; 1.  
 KW Riboflavin biosynthesis; Transferrase; Complete proteome.  
 SQ SEQUENCE 160 AA; 16370 MW; 20EB837C273312EB83 CMC64;

Search completed: February 6, 2003, 11:16:48  
Job time : 7.33333 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: February 6, 2003, 11:16:00 ; Search time 21.3333 Seconds  
(without alignments)  
38.634 Million cell updates/sec

Title: PAT943-4  
Perfect score: 18  
Sequence: 1 slide 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 4129

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 100%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_21:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_protist:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	18	100.0	20	Q49132	Q49132 methylolact
2	18	100.0	33	Q8X4F2	Q8X4F2 escherichia
3	18	100.0	38	Q8S077	Q8S077 bos taurus
4	18	100.0	44	Q8X9P1	Q8X9P1 escherichia
5	18	100.0	45	Q94791	Q94791 tribolium c
6	18	100.0	48	Q8SPU2	Q8SPU2 macaca mula
7	18	100.0	51	Q9NMC6	Q9NMC6 leishmania
8	18	100.0	52	Q941Y6	Q941Y6 oryza sativ
9	18	100.0	53	Q8YL68	Q8YL68 anabaena sp
10	18	100.0	55	Q8ZG82	Q8ZG82 yersinia pe
11	18	100.0	56	Q9Z775	Q9Z775 chlamydia p
12	18	100.0	62	Q9K1Z5	Q9K1Z5 chlamydia p
13	18	100.0	63	Q9Z1S4	Q9Z1S4 pseudomonas
14	18	100.0	63	Q9Z431	Q9Z431 pseudomonas
15	18	100.0	63	Q9R9T3	Q9R9T3 pseudomonas
16	18	100.0	63	Q8RQD2	Q8RQD2 pseudomonas

17	18	100.0	64	10	Q9SJH4	Q9SJH4 arabidopsis
18	18	100.0	65	2	Q9X5B0	Q9X5B0 mycobacteri
19	18	100.0	65	2	Q560Z7	Q560Z7 salmonella
20	18	100.0	65	6	Q9TU12	Q9TU12 oryctolagus
21	18	100.0	66	6	Q9TU11	Q9TU11 sus scrofa
22	18	100.0	66	16	Q06571	Q06571 mycobacteri
23	18	100.0	67	9	Q9MBV1	Q9MBV1 vibrio phag
24	18	100.0	68	13	Q9PMK4	Q9PMK4 esox lucius
25	18	100.0	73	16	Q07227	Q07227 mycobacteri
26	18	100.0	74	5	Q95Y38	Q95Y38 caenorhabdi
27	18	100.0	74	12	Q9DHK2	Q9DHK2 yaba-like d
28	18	100.0	75	2	Q53417	Q53417 bradyrhizob
29	18	100.0	75	10	Q41970	Q41970 arabidopsis
30	18	100.0	75	10	Q9L584	Q9L584 arabidopsis
31	18	100.0	76	16	Q9AK0	Q9AK0 bacillus na
32	18	100.0	76	2	Q9AMZ5	Q9AMZ5 bradyrhizob
33	18	100.0	80	2	Q47775	Q47775 enterococcu
34	18	100.0	80	10	Q81534	Q81534 saccharum o
35	18	100.0	80	16	Q92KC9	Q92KC9 rhizobium m
36	18	100.0	82	10	Q42264	Q42264 arabidopsis
37	18	100.0	84	5	Q9NMB4	Q9NMB4 leishmania
38	18	100.0	84	10	Q94IT9	Q94IT9 fragaria ve
39	18	100.0	84	10	Q94I07	Q94I07 fragaria ve
40	18	100.0	85	2	Q9K65	Q9K65 mycobacteri
41	18	100.0	85	10	Q81399	Q81399 orobanche c
42	18	100.0	87	2	Q93FE2	Q93FE2 helicobacte
43	18	100.0	87	17	Q9HSY5	Q9HSY5 halobacteri
44	18	100.0	88	16	Q8UJ19	Q8UJ19 agrobacteri
45	18	100.0	89	5	Q9NM17	Q9NM17 leishmania

## ALIGNMENTS

RESULT 1	ID	Q49132	PRELIMINARY;	PRT;	20 AA.
AC	Q49132:				
DT	01-NOV-1996 (TREMREL. 01, Created)				
DT	01-NOV-1996 (TREMREL. 01, Last sequence update)				
DT	01-MAR-2002 (TREMREL. 20, Last annotation update)				
DE	Hydroxypyruvate reductase (EC 1.1.1.81) (Fregment).				
GN	hpr.				
OS	Methyllobacterium extorquens.				
OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;				
OC	Methyllobacterium group; Methyllobacterium.				
OX	NCBI_TaxID=408;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=AM1;				
RA	Chistoserdova L., Lidstrom M.;				
RT	"Cloning of a hydroxypyruvate reductase gene from Methyllobacterium				
RT	extorquens AM1, mutagenesis of the gene."				
RL	Submitted (OCT-1991) to the EMBL/Genbank/DBJ databases.				
DR	EMBL: M74226; AAA25377.1; "				
KW	Oxidoreductase; Pyruvate.				
FT	NON_TER				
SQ	SEQUENCE 20 AA: 2325 MW; F50144BAAB7466B CRC64;				
Query Match					
Best Local Similarity	100.0%; Score 18; DB 2; Length 20;				
Matches	4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1 SLDA 4				
DB	12 SLDA 15				
RESULT 2	ID	Q8X4F2	PRELIMINARY;	PRT;	33 AA.
AC	Q8X4F2:				
DT	01-MAR-2002 (TREMREL. 20, Created)				

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
GN 25889.  
OS Escherichia coli O157:H7.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=83334;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN=0157:H7 / EDL933 / ATCC 700927;  
RX MEDLINE=21074935; Pubmed=11206551;  
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
Postal G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
RA Grobeck E.J., Davis N.W., Lam A., Dimlanta E.T., Potamoukis K.,  
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
RA Welch R.A., Blattner F.R.;  
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";  
RL Nature 409:529-533(2001).  
DR EMBL; AE005659; AAG59475.1; -;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 33 AA; 3661 MW; 21166612AB82447B CRC64;

Query Match 100.0%; Score 18; DB 16; Length 33;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SLDA 4  
DB 6 SLDA 9

RESULT 3  
Q8S077 PRELIMINARY; PRT; 38 AA.  
ID 08S077  
AC 08S077;  
DT 01-JUN-2002 (TREMBlrel. 21, Created)  
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Hexokinase I (Fragment).  
GN HK1.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN (1)  
RP SEQUENCE FROM N.A.  
RA Andreoni F., Ruzzo A., Magnani M.;  
RT "Searching for sequence variation in the bovine hexokinase type I  
gene";  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF309639; AAL87545.1; -;  
DR EMBL; AF309638; AAL87545.1; JOINED.  
KW Kinase.  
FT NON\_TER  
FT NON\_TER  
SQ SEQUENCE 38 AA; 4074 MW; 8E931162F96B4F66 CRC64;

Query Match 100.0%; Score 18; DB 6; Length 38;  
Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SLDA 4  
DB 14 SLDA 17  
RESULT 4  
Q8X9P1 PRELIMINARY; PRT; 44 AA.  
AC 08X9P1;

DT 01-MAR-2002 (TREMBlrel. 20, Created)  
DE 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
GN 21184 OR 21623.  
OS Escherichia coli O157:H7.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=83334;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN=0157:H7 / EDL933 / ATCC 700927;  
RX MEDLINE=21074935; Pubmed=11206551;  
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
Postal G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
RA Grobeck E.J., Davis N.W., Lam A., Dimlanta E.T., Potamoukis K.,  
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
RA Welch R.A., Blattner F.R.;  
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";  
RL Nature 409:529-533(2001).  
DR EMBL; AE005311; AAG55738.1; -;  
DR EMBL; AE005274; AAG55329.1; -;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 44 AA; 5079 MW; C640554F24656000 CRC64;

Query Match 100.0%; Score 18; DB 16; Length 44;  
Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SLDA 4  
DB 10 SLDA 13

RESULT 5  
Q94791 PRELIMINARY; PRT; 45 AA.  
ID Q94791  
AC Q94791;  
DT 01-FEB-1997 (TREMBlrel. 02, Created)  
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE TATH2 (Fragment).  
GN TATH2.  
OS Tribolium castaneum (Red flour beetle).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;  
OC Cucujiforma; Tenebrionidae; Tribolium.  
OX NCBI\_TaxID=7070;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97026280; Pubmed=8872459;  
RA Ben-Arie N., McCall A.E., Berkman S., Eichele G., Bellen H.J.,  
RA Zoghbi H.Y.;  
RT "Evolutionary conservation of sequence and expression of the bHLH  
protein Atonal suggests a conserved role in neurogenesis.";  
RL Hum. Mol. Genet. 5:1207-1216(1996).  
DR EMBL; U61152; AAB41307.1; -;  
DR InterPro; IPR001092; HLH\_basic.  
DR Pfam; PF00010; HLH; 1.  
DR SMART; SM00353; HLH; 1.  
FT NON\_TER  
FT NON\_TER  
SQ SEQUENCE 45 AA; 5263 MW; 730A6FE821B82E23 CRC64;

Query Match 100.0%; Score 18; DB 5; Length 45;  
Best Local Similarity 100.0%; Pred. No. 4.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SLDA 4  
DB 26 SLDA 29

RESULT 6  
 08SPU2 PRELIMINARY; PRT: 48 AA.  
 AC 08SPU2;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Caspase 3 (Fragment).  
 GN CASP3.  
 OS Macaca mulatta (Rhesus macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Macaca.  
 NCBI\_TaxID=9544;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RA Norgren R.B. Jr., Zink M.A., Jia Y., Ojeda S.R., Spindel E.R.;  
 RT "Construction of a targeted rhesus macaque microarray."  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF489841; AAL96687.1; -  
 FT NON\_TER 1  
 SO SEQUENCE 48 AA; 5735 MW; D554508750BA06DE CRC64;

Query Match 100.0%; Score 18; DB 6; Length 48;  
 Best Local Similarity 100.0%; Pred. No. 5.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
 ||||  
 Db 22 SLDA 25

RESULT 7  
 09NMC6 PRELIMINARY; PRT: 51 AA.  
 AC 09NMC6;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
 DE Probable kila0944 protein (Fragment).  
 GN LM26.166.  
 OS Leishmania major.  
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
 NCBI\_TaxID=5664;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-FRIEDLIN;  
 RA Murphy L., Quail M., Harris D., Rajandream M., Ivens A., Barrell B.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL160493; CAB97784.1; -  
 FT NON\_TER 51  
 SO SEQUENCE 51 AA; 5755 MW; 4317EFEB839F21872 CRC64;

Query Match 100.0%; Score 18; DB 5; Length 51;  
 Best Local Similarity 100.0%; Pred. No. 5.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
 ||||  
 Db 7 SLDA 10

RESULT 8  
 0941Y6 PRELIMINARY; PRT: 52 AA.  
 AC 0941Y6;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE P0431G06.24 protein (B1139B11.2 protein).  
 GN P0431G06.24 OR B1139B11.2.  
 OS Oryza sativa (Rice), and

OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzoae; Oryza.  
 NCBI\_TaxID=4530, 39947;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. NIPPONBARE;  
 RA Sasaki T., Matsunoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC  
 clone:P0431G06."  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. NIPPONBARE;  
 RA Sasaki T., Matsunoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC  
 clone:B1139B11."  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AP003683; BAB64712.1; -  
 DR EMBL: AP004368; BAB90774.1; -  
 SO SEQUENCE 52 AA; 6048 MW; EA5873D395DC7B33 CRC64;

Query Match 100.0%; Score 18; DB 10; Length 52;  
 Best Local Similarity 100.0%; Pred. No. 5.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
 ||||  
 Db 23 SLDA 26

RESULT 9  
 08YL68 PRELIMINARY; PRT: 53 AA.  
 AC 08YL68;  
 DT 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Hypothetical protein Asr7067.  
 GN ASR7067.  
 OS Anabaena sp. (strain PCC 7120).  
 OC Plasmid PCC7120alpha.  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
 NCBI\_TaxID=103690;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21595285; PubMed=11759840;  
 RA Kareko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
 RA Watanabe A., Iriyuchi M., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kohara M., Matsunoto M., Matsuno A., Muraki A.,  
 RA Nakazaki N., Shimo S., Sugimoto M., Takasawa M., Yamada M.,  
 RA Yasuda M., Tabata S.;  
 RT "Complete genomic sequence of the filamentous nitrogen-fixing  
 cyanobacterium Anabaena sp. strain PCC 7120."  
 RL DNA Res. 8:205-213(2001).  
 DR EMBL: AP003600; BAB78151.1; -  
 KW Plasmid; Hypothetical protein; Complete proteome.  
 SO SEQUENCE 53 AA; 5639 MW; 3D75D4A82284175 CRC64;

Query Match 100.0%; Score 18; DB 16; Length 53;  
 Best Local Similarity 100.0%; Pred. No. 5.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
 ||||  
 Db 43 SLDA 46

RESULT 10  
 08ZG82 PRELIMINARY; PRT: 55 AA.  
 AC 08ZG82;  
 ID 08ZG82;

DT 01-MAR-2002 (TREMblrel. 20, Created)  
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
 DE Putative ribosome modulation factor.  
 GN RME OR YP04423A.  
 OS Yersinia pestis.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Yersinia.  
 OX NCB1\_Taxid=632;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CO-92 / BIOVAR ORIENTALIS;  
 RX MEDLINE=21470413; PubMed=11586360;  
 RA Parkhill J., Wren B.W., Thomson N.R., James K.D., Churcher C., Mungall K.L.,  
 RA Prentice M.B., Sebatia M., James N.R., Churcher C., Mungall K.L.,  
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,  
 RA Feltham T., Hamlin R., Holtroyd S., Jagsels K., Karlyshev A.V.,  
 RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,  
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrett B.G.;  
 RT "Genome sequence of Yersinia pestis, the causative agent of plague."  
 RL Nature 413:523-527(2001).  
 DR EMBL: AJ141418; CAC30253.1; -  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 55 AA; 6406 MW; B7CE6B7FEFDAD6E2 CRC64;

Query Match 100.0%; Score 18; DB 16; Length 55;  
 Best Local Similarity 100.0%; Pred. No. 6.1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
 Db 33 SLDA 36

RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,  
 RA Gwin M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,  
 RA Eissen J., Fraser C.M.;  
 RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia  
 RT Pneumoniae AR39."  
 RL Nucleic Acids Res. 28:1397-1406(2000).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=J138;  
 RX MEDLINE=20330349; PubMed=10871362;  
 RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,  
 RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;  
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138  
 RT from Japan and CML029 from USA."  
 RL Nucleic Acids Res. 28:2311-2314(2000).  
 DR EMBL: AE001663; AAD18968.1; -  
 DR EMBL: AB035941; BAA88649.1; -  
 DR EMBL: AP002548; BAA99039.1; -  
 DR EMBL: AE002261; AAF38814.1; -  
 DR TIGR: CP1039; -  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 56 AA; 6344 MW; 057543635D28B661 CRC64;

Query Match 100.0%; Score 18; DB 16; Length 56;  
 Best Local Similarity 100.0%; Pred. No. 6.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
 Db 47 SLDA 50

RESULT 11  
 ID 092775 PRELIMINARY; PRT; 56 AA.  
 AC 092775;  
 DT 01-MAY-1999 (TREMblrel. 10, Created)  
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)  
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
 DE Hypothetical protein CPN0831.  
 GN CPN0831 OR CPJ0831 OR CPJ039.  
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCB1\_Taxid=83558;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CWL029;  
 RX MEDLINE=99206606; PubMed=10192388;  
 RA Kalman S., Mitchell W., Marathe R., Lamme C., Fan J., Hyman R.W.,  
 RA Olinger U., Grimwood J., Davis R.W., Stephens R.S.;  
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis."  
 RT Nat. Genet. 21:385-389(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=J138;  
 RA Hirakawa H., Shirai M., Kuhara S.;  
 RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=J138;  
 RA Shirai M.;  
 RT "Genomic sequence comparison of two unrelated isolates of chlamydia  
 RT pneumoniae from Japan and U.S.";  
 RT Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AR39;  
 RX MEDLINE=20150255; PubMed=10684935;  
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,  
 RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,  
 RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,

RESULT 12  
 ID 09K125 PRELIMINARY; PRT; 62 AA.  
 AC 09K125;  
 DT 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)  
 DE Hypothetical protein CP0756.  
 GN CP0756.  
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCB1\_Taxid=83558;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AR39;  
 RX MEDLINE=20150255; PubMed=10684935;  
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,  
 RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,  
 RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,  
 RA Gwin M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,  
 RA Eissen J., Fraser C.M.;  
 RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia  
 RT Pneumoniae AR39."  
 RL Nucleic Acids Res. 28:1397-1406(2000).  
 DR EMBL: AE002235; AAF38560.1; -  
 DR TIGR: CP0756; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 62 AA; 6840 MW; C07FE52E269FE9E CRC64;

Query Match 100.0%; Score 18; DB 16; Length 62;  
 Best Local Similarity 100.0%; Pred. No. 6.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
 Db 32 SLDA 35

RESULT 13  
 ID 092154 PRELIMINARY; PRT; 63 AA.



AC 092154;  
DT 01-MAY-1999 (TReMBLrel. 10, Created)  
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
DE 4-oxalocrotonate isomerase.  
GN NAHU.  
OS Pseudomonas stutzeri (Pseudomonas perfectomarina).  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=316;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AN10;  
RX MEDLINE=20179685; PubMed=10713446;  
RA Bosch R., Garcia-Valdes E., Moore E.R.B.;  
RT "Complete nucleotide sequence and evolutionary significance of a  
RT chromosomally encoded naphthalene-degradation lower pathway from  
RT Pseudomonas stutzeri AN10.";  
RL Gene 245:65-74(2000).  
DR EMBL: AF039534; AAD02155.1; -.  
DR HSSP: Q01468; 1BJP.  
DR InterPro: IPR004370; Taut.  
DR Pfam: PF01361; Tautomerase; 1.  
DR ProDom: PD019232; Tautomerase; 1.  
DR TIGRFAMs: TIGR00013; taut; 1.  
KW Isomerase.  
SQ SEQUENCE 63 AA; 6883 MW; 3FF05827BC926F87 CRC64;

Query Match 100.0%; Score 18; DB 2; Length 63;  
Best Local Similarity 100.0%; Pred. No. 7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
Db 31 SLDA 34

RESULT 14  
ID 092431 PRELIMINARY; PRT; 63 AA.  
AC 092431;  
DT 01-MAY-1999 (TReMBLrel. 10, Created)  
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
DE 4-oxalocrotonate tautomerase.  
GN NAHU.  
OS Pseudomonas putida.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=303;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=67;  
RX MEDLINE=99255564; PubMed=10322041;  
RA Grimm A.C., Harwood C.S.;  
RT "Nah<sub>1</sub>, a catabolic plasmid-encoded receptor required for chemotaxis of  
RT Pseudomonas putida to the aromatic hydrocarbon naphthalene.";  
RL J. Bacteriol. 181:3310-3316(1999).  
DR EMBL: AF100302; AAD13221.1; -.  
DR HSSP: P49172; 1OTF.  
DR InterPro: IPR004370; Taut.  
DR Pfam: PF01361; Tautomerase; 1.  
DR ProDom: PD019232; Tautomerase; 1.  
DR TIGRFAMs: TIGR00013; taut; 1.  
KW Plasmid.  
SQ SEQUENCE 63 AA; 7122 MW; 2EE96A4BA328B32F CRC64;

Query Match 100.0%; Score 18; DB 2; Length 63;  
Best Local Similarity 100.0%; Pred. No. 7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4

Db 31 SLDA 34  
RESULT 15  
ID 09R9T3 PRELIMINARY; PRT; 63 AA.  
AC 09R9T3;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
DE 4-oxalocrotonate tautomerase.  
GN NAHU.  
OS Pseudomonas putida.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=303;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PI;  
RA van Beilen J.B., Roehlisberger M., Witholt B.;  
RT "Organization of the Pseudomonas putida (oleovorans) Gp01 and  
RT Pseudomonas putida Pl alkane oxidation genes.";  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ249793; CAB66148.1; -.  
DR HSSP: P49172; 1OTF.  
DR InterPro: IPR004370; Taut.  
DR Pfam: PF01361; Tautomerase; 1.  
DR ProDom: PD019232; Tautomerase; 1.  
DR TIGRFAMs: TIGR00013; taut; 1.  
SQ SEQUENCE 63 AA; 6996 MW; A16F7F28E9C91AF3 CRC64;

Query Match 100.0%; Score 18; DB 2; Length 63;  
Best Local Similarity 100.0%; Pred. No. 7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
Db 31 SLDA 34

RESULT 16  
ID 08R0D2 PRELIMINARY; PRT; 63 AA.  
AC 08R0D2;  
DT 01-JUN-2002 (TReMBLrel. 21, Created)  
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
DE 4-olalocrotonate tautomerase.  
GN XYIH.  
OS Pseudomonas putida.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=303;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PAW630;  
RA Wessler H.G., Hares D.R., Poulter M.D., Voss J.A., Khedairy H.S.,  
RA Baker R.F., Azadpour E.E., Luo X., Benjamin R.C.;  
RT "Sequence of the lower (meta-cleavage) xyl operon of the Pseudomonas  
RT putida TOL plasmid pPK1.";  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF134348; AAL83667.1; -.  
KW Plasmid.  
SQ SEQUENCE 63 AA; 6969 MW; 23805E3F916B5802 CRC64;

Query Match 100.0%; Score 18; DB 2; Length 63;  
Best Local Similarity 100.0%; Pred. No. 7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4

Db 31 SLDA 34

RESULT 17

Q9SJH4 PRELIMINARY; PRT; 64 AA.

AC Q9SJH4 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)

DE AT2G42820 protein.

GN AT2G42820

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

OC NCBI\_TaxID=3702;

OX NCBI\_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV, COLUMBIA;

RX MEDLINE=20083487; PubMed=10617197;

RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D., Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V., Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S., Adams M.D., Shenn M., Vanden S.E., Umayam L., Tallon L.J., Gill J.E., Cronin L.A., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R., Coppenhaver G.P., Preuss D., Nieman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M., Venter J.C.;

RA "Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana."

RT thaliana."

RL Nature 402:761-768(1999).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=CV, COLUMBIA;

RA Lin X.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AC006931; AAD21726.1; -

SO SEQUENCE 64 AA; 7401 MW; A8BD423F5B099C0 CRC64;

Query Match

Best Local Similarity 100.0%; Score 18; DB 10; Length 64;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4

Db 33 SLDA 36

RESULT 18

Q9X5B0 PRELIMINARY; PRT; 65 AA.

AC Q9X5B0 01-NOV-1999 (TREMBlrel. 12, Created)

DT 01-NOV-1999 (TREMBlrel. 12, last sequence update)

DE DnaJ protein (Fragment).

OS Mycobacterium xenopi.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.

OC NCBI\_TaxID=1789;

OX NCBI\_TaxID=1789;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=3945;

RA Brown T.J.;

RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF123349; AAD30982.1; -

DT HSP; P25685; 1HDJ.

DE Interpro: IPR001623; DnaJ\_N.

DR Interpro: IPR003095; Hsp\_DnaJ.

PLam: PF00226; DnaJ; 1.

DR PRINTS: PR00625; DnaJPROTEIN.

DR SMART: SM00271; DnaJ; 1.

DR PROSITE: PSS0076; DnaJ\_2; 1.

FT NON\_TER 1 1

FT NON\_TER 65 65

SO SEQUENCE 65 AA; 7486 MW; 92BE516FDE2CB471 CRC64;

Query Match

Best Local Similarity 100.0%; Score 18; DB 2; Length 65;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4

Db 15 SLDA 18

RESULT 19

Q56027 PRELIMINARY; PRT; 65 AA.

AC Q56027 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)

DE S9PA (Fragment).

GN S9PA.

OS Salmonella typhimurium.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Salmonella.

OC NCBI\_TaxID=602;

OX NCBI\_TaxID=602;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SL1344;

RX MEDLINE=96342378; PubMed=8748032;

RA Hueck C.J., Hantman M.J., Bajaj V., Johnston C., Lee C.A., Miller S.I.;

RA "Salmonella typhimurium secreted invasion determinants are homologous to Shigella ipa proteins."

RT Mol. Microbiol. 18:479-490(1995).

RL EMBL: U30491; AAC43549.1; -

FT NON\_TER 65 65

SO SEQUENCE 65 AA; 6839 MW; 86269E85D08E4CB5 CRC64;

Query Match

Best Local Similarity 100.0%; Score 18; DB 2; Length 65;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4

Db 62 SLDA 65

RESULT 20

Q9TU12 PRELIMINARY; PRT; 65 AA.

AC Q9TU12 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)

DE Uroplakin 3 (Fragment).

OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

OC NCBI\_TaxID=9986;

OX NCBI\_TaxID=9986;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NEW ZEALAND WHITE; TISSUE=BLADDER UROTHELIUM;

RA Cho S.K., Kwon D.N., Chu H.J., Kim J.H.;

RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF178938; AAD55729.1; -

FT NON\_TER 65 65

SO SEQUENCE 65 AA; 7107 MW; DC18F18C18EA6362 CRC64;

Query Match

Best Local Similarity 100.0%; Score 18; DB 6; Length 65;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
1111

DB 11 SLDA 14

RESULT 21

Q9TUL1 PRELIMINARY; PRT; 66 AA.

AC Q9TUL1; 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)

OS Uroplakin 3 (Fragment).

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.

OC NCBI\_Taxid=9823;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=DURC X LANDRACE X YORKSHIRE; TISSUE=BLADDER UROTHELIUM;

RA Kwon D.N., Chu H.J., Cho S.K., Kim J.H.;

RT "Cloning of pig uroplakin III cDNA."

RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF178939; AAD55730.1; -

FT NON\_TER 1 1

FT SEQUENCE 66 AA; 7190 MW; 286DF9159818E36A CRC64;

Query Match Best Local Similarity 100.0%; Score 18; DB 6; Length 66; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
1111

DB 11 SLDA 14

RESULT 22

Q06571 PRELIMINARY; PRT; 66 AA.

AC Q06571; 01-JUL-1997 (TREMBLrel. 04, Created)

DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

OS Hypothetical protein RV1119C.

GN RV1119C OR MT1151 OR MTCY2268.08C.

OS Mycobacterium tuberculosis.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacteriaceae; Mycobacterium.

OC NCBI\_Taxid=1773;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=H37RV;

RA MEDLINE=98295987; PubMed=9634230;

RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekoa F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sultson J.E., Taylor K., Whitehead S., Barrett B.G.;

RA "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence."

RL Nature 393:537-544(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=CDC 1551 / OSHKOSH;

RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J., Deboy R., Dodson R., Gwinn M., Haft D., Hickey E., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,

RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A., Bishai W.;

RT "Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains."

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: Z95585; CAB09051.1; ALT\_INIT.

DR EMBL: AE006994; AAK45408.1; -

DR TIGR: MT1151; -

DR Tubercultist; RV1119C; -

KW Hypothetical protein; Complete proteome.

SEQUENCE 66 AA; 6929 MW; 582B7AD3C58CF10 CRC64;

Query Match Best Local Similarity 100.0%; Score 18; DB 16; Length 66; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
1111

DB 56 SLDA 59

RESULT 23

Q9MBV1 PRELIMINARY; PRT; 67 AA.

AC Q9MBV1; 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

OS Repressor-like protein.

GN RSTR.

OS Vibrio phage CTX.

OC Viruses; ssDNA viruses; Inoviridae; Inovirus.

OC NCBI\_Taxid=141904;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=21359315; PubMed=11466276;

RA Mukhopadhyay A.K., Chakraborty S., Takeda Y., Nair G.B., Berg D.E.;

RT "Characterization of VPI Pathogenicity Island and CTXphi Prophage in Environmental Strains of Vibrio cholerae."

RL J. Bacteriol. 183:4737-4746(2001).

DR EMBL: AF13308; AAF43267.1; -

SO SEQUENCE 67 AA; 7586 MW; 5EB044E3840C0693 CRC64;

Query Match Best Local Similarity 100.0%; Score 18; DB 9; Length 67; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
1111

DB 7 SLDA 10

RESULT 24

Q9PWK4 PRELIMINARY; PRT; 68 AA.

AC Q9PWK4; 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

OS Arylsulfatase N-acetyltransferase (EC 2.3.1.87) (Fragment).

GN AANAT1.

OS Esoc lucius (Northern pike).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Esociformes; Esocidae; Esoc.

OC NCBI\_Taxid=8010;

RN [1]

RP SEQUENCE FROM N.A.

RC MEDLINE=99185141; PubMed=10085157;

RA Coon S.L., Begay V., Deurloo D., Falcon J., Klein D.C.;

RT "Two arylalkylamine N-acetyltransferase genes mediate melanin synthesis in fish."

RL J. Biol. Chem. 274:9076-9082(1999).

DR EMBL; AF034084; AAD21319.1; -.  
 DR HSP; Q29495; ICJW.  
 KW Acyltransferase; Transferase.  
 FT NON\_TER 1  
 FT NON\_TER 68  
 SO SEQUENCE 68 AA; 7711 MW; 38C000CF44BB5C48 CRC64;

Query Match 100.0%; Score 18; DB 13; Length 68;  
 Best Local Similarity 100.0%; Pred. No. 7.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
 ||||  
 DB 59 SLDA 62

## RESULT 25

ID 007227 PRELIMINARY; PRT; 73 AA.

AC 007227;

DT 01-JUL-1997 (TREMBLrel. 04, Created)

DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE Hypothetical protein Rv0300.

GN Rv0300 OR MTC163.05.

OS Mycobacterium tuberculosis.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.

OX NCBI\_TaxID=1773;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=H37RV;

RA MEDLINE=98295987; PubMed=9634230;

RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,

RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,

RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,

RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holtroyd S.,

RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,

RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,

RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,

RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;

RT "Deciphering the biology of Mycobacterium tuberculosis from the

RT complete genome sequence."

RL Nature 393:537-544(1998).

EMBL; Z96800; CAB09591.1; -.

DR Tuberculolist; Rv0300; -.

DR InterPro; IPR002145; COPG\_HTH\_4.

DR Pfam; PF01402; HTH\_4; 1.

KW Hypothetical protein; Complete proteome.

SO SEQUENCE 73 AA; 8088 MW; 379D16E6EA7D7979 CRC64;

Query Match 100.0%; Score 18; DB 16; Length 73;

Best Local Similarity 100.0%; Pred. No. 8.3e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
 ||||

DB 16 SLDA 19

## RESULT 26

ID 095Y38 PRELIMINARY; PRT; 74 AA.

AC 095Y38;

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE Hypothetical 8.3 kDa protein.

GN Y77E11A.14

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Pelodierinae; Caenorhabditis.

OX NCBI\_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BRISTOL N2;

RX MEDLINE=99069613; PubMed=9851916;

RA None;

RT "Genome sequence of the nematode C. elegans: a platform for

RT investigating biology. The C. elegans Sequencing Consortium."

RL Science 282:2012-2018(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=BRISTOL N2;

RA Dante M.;

RT "The sequence of C. elegans cosmid Y77E11A."

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=BRISTOL N2;

RA Waterston R.;

RT "Direct Submission."

RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC024214; AAK67241.1; -.

KW Hypothetical protein.

SO SEQUENCE 74 AA; 8307 MW; BEAAC287BF389C3 CRC64;

Query Match 100.0%; Score 18; DB 5; Length 74;

Best Local Similarity 100.0%; Pred. No. 8.4e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
 ||||

DB 60 SLDA 63

## RESULT 27

ID 09DHK2

AC 09DHK2; PRELIMINARY; PRT; 74 AA.

DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE 111L protein.

GN 111L.

OS Yaba-like disease virus (YLDV).

OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;

OC Yatapoxvirus.

OX NCBI\_TaxID=132475;

RN [1]

RP SEQUENCE FROM N.A.

RC Lee H.J.;

RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=21176366; PubMed=11277691;

RA Lee H.J., Essani K., Smith G.L.;

RT "The genome sequence of Yaba-like disease virus, a yatapoxvirus."

RL Virology 281:170-192(2001).

RN [3]

RP SEQUENCE FROM N.A.

RA Lee H.J.;

RT Thesis (2000), Sir William Dunn School of Pathology, University of.

DR EMBL; AJ293568; CAC21349.1; -.

SO SEQUENCE 74 AA; 7905 MW; 70B1546003F00823 CRC64;

Query Match 100.0%; Score 18; DB 12; Length 74;

Best Local Similarity 100.0%; Pred. No. 8.4e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
 ||||

DB 44 SLDA 47

## RESULT 28

```

053417
ID 053417      PRELIMINARY;      PRT:      75 AA.
AC 053417;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE NOLZ.
GN NOLZ.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bradyrhizobium group; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94281668; PubMed=8012039;
RA Dockendorf T.C., Sharma A.J., Stacey G.;
RT "Identification and characterization of the nolyz genes of
RT Bradyrhizobium japonicum."
RL Mcl. Plant Microbe Interact. 7:173-180(1994).
DR EMBL: S70736; AAB31118.2; -.
SQ SEQUENCE 75 AA; 8662 MW; 666068FAC2684CDA CRC64;

Query Match      100.0%; Score 18; DB 2; Length 75;
Best Local Similarity 100.0%; Pred. NO. 8.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4
DB 45 SLDA 48

RESULT 29
041970
ID 041970      PRELIMINARY;      PRT:      75 AA.
AC 041970;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ADP ATP carrier protein (Fragment).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CY. COLOMBIA; TISSUE=SEEDLING;
RA Desprez T., Amselem J., Chapello H., Rouze P., Caboche M., Hofte H.;
RL Submitted (NOV-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z18781; CAA79257.1; -.
DR InterPro: IPR001993; Mitoch_carrier.
DR Pfam: PF00153; mito_carr; 1.
FT NON_TER
SQ SEQUENCE 75 AA; 7913 MW; 7406BA4881A994B2 CRC64;

Query Match      100.0%; Score 18; DB 10; Length 75;
Best Local Similarity 100.0%; Pred. NO. 8.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4
DB 21 SLDA 24

RESULT 30
091J84
ID 091J84      PRELIMINARY;      PRT:      75 AA.
AC 091J84;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Genomic DNA, chromosome 3, BAC clone:F21A17.
OS Arabidopsis thaliana (Mouse-ear cress).

```

```

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20363099; PubMed=10907853;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the regions of 4,251,695 bp covered by ninety pl,
RT TAC and BAC clones."
RL DNA Res. 7:217-221(2000).
DR EMBL: AP000732; BAB01204.1; -.
SQ SEQUENCE 75 AA; 8777 MW; 5D23536B9B64143A CRC64;

Query Match      100.0%; Score 18; DB 10; Length 75;
Best Local Similarity 100.0%; Pred. NO. 8.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4
DB 2 SLDA 5

RESULT 31
09KAKO
ID 09KAKO      PRELIMINARY;      PRT:      76 AA.
AC 09KAKO;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE Thioedoxin related protein.
GN BH2287.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis."
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL: AP001515; BAB06006.1; -.
KW Complete Proteome.
SQ SEQUENCE 76 AA; 8916 MW; 6028CDFC766C6307 CRC64;

Query Match      100.0%; Score 18; DB 16; Length 76;
Best Local Similarity 100.0%; Pred. NO. 8.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4
DB 55 SLDA 58

RESULT 32
09AMZ5
ID 09AMZ5      PRELIMINARY;      PRT:      78 AA.
AC 09AMZ5;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)

```

DE NOLZ.  
GN NOLZ.  
OS Bradyrhizobium japonicum.  
OC Bacteria: Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Bradyrhizobium group; Bradyrhizobium.  
OX NCBI\_TaxID=375;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=110SPC4;  
RX MEDLINE=21101824; PubMed=11157954;  
RA Gottfert M., Rochlisberger S., Kundig C., Beck C., Marty R.,  
RA Hennecke H.;  
RT "Potential symbiosis-specific genes uncovered by sequencing a 410-kb  
RT DNA region of the Bradyrhizobium japonicum chromosome.";  
RL J. Bacteriol. 183:1405-1412(2001).  
DR EMBL: AF322013; AAC60986.1;  
SQ SEQUENCE 78 AA; 8994 MW; 8AA5AAFF174C86355 CRC64;

Query Match 100.0%; Score 18; DB 2; Length 78;  
Best Local Similarity 100.0%; Pred. No. 8.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
DB 48 SLDA 51

RESULT 33  
O47775  
ID 047775 PRELIMINARY; PRT; 80 AA.  
AC 047775;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE ORF9.  
GN ORF9.  
OS Enterococcus faecalis (Streptococcus faecalis).  
OC Plasmid pYL17.  
OC Bacteria: Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
OC Enterococcaceae; Enterococcus.  
OX NCBI\_TaxID=1351;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Tomita H.;  
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL: D78257; BA11326.1;  
KW Plasmid.  
SQ SEQUENCE 80 AA; 9100 MW; 4F1F97DF1500C169 CRC64;

Query Match 100.0%; Score 18; DB 2; Length 80;  
Best Local Similarity 100.0%; Pred. No. 9.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
DB 31 SLDA 34

RESULT 34  
O81534  
ID 081534 PRELIMINARY; PRT; 80 AA.  
AC 081534;  
DT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Chalcone synthase (Fragment).  
GN CHS.  
OS Saccharum officinarum (Sugarcane).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
OC Panicoideae; Andropogoneae; Saccharum.  
OX NCBI\_TaxID=4547;  
RN [1]

RP SEQUENCE FROM N.A.  
RA Veronesi C., Thailouarn P.;  
RT "Sugarcane resistance to Striga hermonthica.";  
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF079174; AAC27797.1;  
DR InterPro: IPR001099; N-C\_synthase.  
DR Pfam: PF00195; Chal\_syl\_synth; 1.  
DR Prodom: PD000453; N-C\_synthase; 1.  
FT NON\_TER 1  
FT NON\_TER 80  
SQ SEQUENCE 80 AA; 8744 MW; 7B94776D2A46F418 CRC64;

Query Match 100.0%; Score 18; DB 10; Length 80;  
Best Local Similarity 100.0%; Pred. No. 9.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
DB 12 SLDA 15

RESULT 35  
Q92KC9  
ID Q92KC9 PRELIMINARY; PRT; 80 AA.  
AC Q92KC9;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Hypothetical protein R01596.  
GN R01596 OR SMC01197.  
OS Rhizobium meliloti (Sinorhizobium meliloti).  
OC Bacteria: Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Sinorhizobium.  
OX NCBI\_TaxID=382;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=1021;  
RX MEDLINE=21396507; PubMed=11481430;  
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,  
RA Bolstead P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,  
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaie V., Masuy D.,  
RA Pohl T., Portetlelle D., Puehler A., Purnelle B., Ramsperger U.,  
RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galbert F.;  
RT "Analysis of the chromosome sequence of the legume symbiont  
RT Sinorhizobium meliloti strain 1021.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
DR EMBL: AU591787; CAC46175.1;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 80 AA; 9461 MW; 7C02CDSF7830B99D CRC64;

Query Match 100.0%; Score 18; DB 16; Length 80;  
Best Local Similarity 100.0%; Pred. No. 9.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
DB 37 SLDA 40

RESULT 36  
Q42264  
ID Q42264 PRELIMINARY; PRT; 82 AA.  
AC Q42264;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Carrier protein (Fragment).  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN-CV. COLUMBIA; TISSUE=SEEDLING;  
RA Desprez T., Amsellem J., Chiapello H., Caboche M., Hofte H.;  
RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Z33799; CA83932.1; -  
DR InterPro: IPR00193; Mitoch\_carrier.  
DR Pfam: PF00153; mito\_carr; 1.  
FT NON\_TER 1  
SQ SEQUENCE 82 AA; 8559 MW; 619712734ECABE12 CRC64;

Query Match 100.0%; Score 18; DB 10; Length 82;  
Best Local Similarity 100.0%; Pred. No. 9.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|  
|  
|  
|  
Db 28 SLDA 31

## RESULT 37

ID Q9NMB4 PRELIMINARY; PRT; 84 AA.  
AC Q9NMB4;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Possible CGI-17 protein (Fragment).  
GN LM26.178.  
OS Leishmania major.  
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
ON NCBI\_TaxID=5664;  
RX [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-FRIEDLIN;  
RA Murphy L., Quail M., Harris D., Rajandream M., Ivens A., Barrell B.;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL160493; CAB97796.1; -  
DR InterPro: IPR005141; eRF1\_2.  
DR Pfam: PF03464; eRF1\_2; 1.  
FT NON\_TER 84  
SQ SEQUENCE 84 AA; 9627 MW; 9240E0AA14701A8F CRC64;

Query Match 100.0%; Score 18; DB 5; Length 84;  
Best Local Similarity 100.0%; Pred. No. 9.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|  
|  
|  
|  
Db 13 SLDA 16

## RESULT 38

ID Q94IT9 PRELIMINARY; PRT; 84 AA.  
AC Q94IT9;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Chalcone synthase (Fragment).  
OS Fragaria vesca subsp. bracteata.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids I; Rosales; Rosaceae; Rosoideae; Fragaria.  
ON NCBI\_TaxID=157670;  
RX [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. DNIC;  
RA Deng C., Davis T.;  
RT "Molecular identification of the yellow fruit color (c) locus in  
theor. Appl. Genet. 0:0-0(2001)."  
RL EMBL; AY017485; AAK33142.1; -  
DR InterPro: IPR001099; N-C\_synthase.

DR Pfam: PF00195; Chal\_stil\_synth; 1.  
DR ProDom: PD000453; N-C\_synthase; 1.  
FT NON\_TER 1  
SQ SEQUENCE 84 AA; 10084 MW; FE2F87B3AF485739 CRC64;

Query Match 100.0%; Score 18; DB 10; Length 84;  
Best Local Similarity 100.0%; Pred. No. 9.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|  
|  
|  
|  
Db 55 SLDA 58

## RESULT 39

ID Q94IU7 PRELIMINARY; PRT; 84 AA.  
AC Q94IU7;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Chalcone synthase (Fragment).  
OS Fragaria vesca subsp. vesca.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids I; Rosales; Rosaceae; Rosoideae; Fragaria.  
ON NCBI\_TaxID=101020;  
RX [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. YELLOW WONDER;  
RA Deng C., Davis T.;  
RT "Molecular identification of the yellow fruit color (c) locus in  
theor. Appl. Genet. 0:0-0(2001)."  
RL EMBL; AY017477; AAK33134.1; -  
DR InterPro: IPR001099; N-C\_synthase.  
DR Pfam: PF00195; Chal\_stil\_synth; 1.  
DR ProDom: PD000453; N-C\_synthase; 1.  
FT NON\_TER 84  
SQ SEQUENCE 84 AA; 10015 MW; FE2EC3F6AE0C1239 CRC64;

Query Match 100.0%; Score 18; DB 10; Length 84;  
Best Local Similarity 100.0%; Pred. No. 9.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
|  
|  
|  
|  
Db 55 SLDA 58

## RESULT 40

ID Q9KK65 PRELIMINARY; PRT; 85 AA.  
AC Q9KK65;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE Exported protein 996A160 (Fragment).  
OS Mycobacterium avium.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
ON NCBI\_TaxID=1764;  
RX [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-950A1;  
RA Carroll J.D., Wallace R.C., Keane J., Arbelt R.D.;  
RT "Identification of Mycobacterium avium DNA Sequences that Encode  
Exported Proteins by using phoA Gene fusions."  
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF143401; AAF74995.1; -  
FT NON\_TER 1

FT NON\_TER 85 85  
SQ SEQUENCE 85 AA; 9181 MW; FAF1A39B1F6C7944 CRC64;

Query Match  
Best Local Similarity 100.0%; Score 18; DB 2; Length 85;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
1111  
Db 53 SLDA 56

## RESULT 41

ID 081399 PRELIMINARY; PRT; 85 AA.  
AC 081399;  
DT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
GN Chalcone synthase (Fragment).  
OS Orobanchaceae.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; easterids I; Lamiales; Orobanchaceae; Orobanchae.  
OX NCBI\_TaxID=78542;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Veronesi C., Labrousse P., Thalonarn P.;  
RT "Histological, physiological and molecular aspects of sunflower  
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF074401; AAC26132.1; -  
DR InterPro; IPR001099; N-C-synthase.  
DR Pfam; PF00195; Chalcone synthase; 1.  
DR ProDom; PD000453; N-C-synthase; 1.  
FT NON\_TER 1  
FT NON\_TER 85  
SQ SEQUENCE 85 AA; 9483 MW; B9A6ABEDD6D4777E CRC64;

Query Match  
Best Local Similarity 100.0%; Score 18; DB 10; Length 85;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
1111  
Db 12 SLDA 15

## RESULT 42

ID 093EF2 PRELIMINARY; PRT; 87 AA.  
AC 093EF2;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
GN Hypothetical 10.3 kDa protein (Fragment).  
OS Helicobacter hepaticus.  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=32025;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=3B1;  
RA Ge Z., Feng Y., Fox J.G.;  
RT "Helicobacter hepaticus genome: construction of an ordered cosmid  
RL library and sequence analysis of the selected genomic regions.";  
DR EMBL; AF358712; AAL16703.1; -  
GN Hypothetical protein.  
FT NON\_TER 1  
FT NON\_TER 87  
SQ SEQUENCE 87 AA; 10264 MW; 406A612F41A3EFD8 CRC64;

Query Match  
Best Local Similarity 100.0%; Score 18; DB 2; Length 87;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
1111  
Db 24 SLDA 27

## RESULT 43

ID 09HSY5 PRELIMINARY; PRT; 87 AA.  
AC 09HSY5;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DE Vng0025H.  
GN Vng0025H.  
OS Halobacterium sp. (strain NRC-1).  
OC Archaea; Euryarchaeota; Halobacteriales;  
OC Halobacteriaceae; Halobacterium.  
OX NCBI\_TaxID=64091;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20504483; PubMed=11016950;  
RA Ng W.V., Kennedy S.P., Maniatis G.G., Bergquist B., Pan M.,  
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Shrogha J.,  
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welt R., Goo Y.A.,  
RA Lelthausen B., Keller K., Cruz R., Danon M.J., Hough D.W.,  
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,  
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudis J.L., Jung K.-H.,  
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
RA Eshart H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;  
RT "Genome sequence of Halobacterium species NRC-1.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
DR EMBL; AE004971; AAC18665.1; -  
GN Complete proteome.  
SQ SEQUENCE 87 AA; 10001 MW; 59334C62386E58F CRC64;

Query Match  
Best Local Similarity 100.0%; Score 18; DB 17; Length 87;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLDA 4  
1111  
Db 9 SLDA 12

## RESULT 44

ID 08UJ19 PRELIMINARY; PRT; 88 AA.  
AC 08UJ19;  
DT 01-JUN-2002 (TREMBLrel. 21, Created)  
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
GN Hypothetical protein Atu0119.  
GN Atu0119 OR AGR.C.187.  
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Rhizobium.  
OX NCBI\_TaxID=176299;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21608550; PubMed=11743193;  
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,  
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,  
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boye D., Sr.,  
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,  
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,  
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,  
RA Zhang S., Too H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,  
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,



RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,  
 RA Nester E.W.,  
 RT "The genome of the natural genetic engineer *Agrobacterium tumefaciens*  
 RT C58."  
 RL Science 294:2317-2323(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-21608551; PubMed-11743194;  
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,  
 RA Gurrillo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,  
 RA Houmlet K., Gordon J., Vaudin M., Tatchouk O., Epp A., Liu F.,  
 RA William C., Allinger M., Doughy D., Scott C., Lappas C., Markelz B.,  
 RA Flanagan C., Crowell C., Gurson J., Lomo C., Seear C., Strub G.,  
 RA Cielo C., Slater S.,  
 RT "Genome sequence of the plant pathogen and biotechnology agent  
 RT *Agrobacterium tumefaciens* C58."  
 RL Science 294:2323-2328(2001).  
 DR EMBL; AE008986; AAL414.1; -;  
 DR EMBL; AE007953; AAK85939.1; -;  
 KM Hypothetical protein: Complete proteome.  
 SQ SEQUENCE 88 AA; 10281 MW; 1E2F03740FF126F8 CRC64;

Query Match 100.0%; Score 18; DB 16; Length 88;  
 Best Local Similarity 100.0%; Pred. No. 1e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
 ||||  
 DB 76 SLDA 79

## RESULT 45

O9NM17 PRELIMINARY; PRT; 89 AA.  
 AC O9NM17;  
 DT 01-OCT-2000 (TRENBLREL. 15, Created)  
 DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)  
 DT 01-OCT-2000 (TRENBLREL. 15, Last annotation update)  
 DE Hypothetical 9.8 kda protein (Fragment).  
 GN LM26.105.  
 OS Leishmania major.  
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
 OX NCBI\_TaxID=5664;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-FRIEDLIN;  
 RA Murphy L., Quail M., Harris D., Rajandream M., Ivens A., Barrell B.,  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL160493; CAB9723.1; -;  
 KM Hypothetical protein.  
 FT NON\_PEP 89  
 SQ SEQUENCE 89 AA; 9751 MW; BD4B82F20955A000 CRC64;

Query Match 100.0%; Score 18; DB 5; Length 89;  
 Best Local Similarity 100.0%; Pred. No. 1e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLDA 4  
 ||||  
 DB 13 SLDA 16

Search completed: February 6, 2003, 11:22:00  
 Job time : 23.3333 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 11:16:00 : Search time 27.8333 Seconds  
(Without alignments)  
19.150 Million cell updates/sec

Title: PAT943-5  
Perfect score: 23  
Sequence: 1 nyda 4

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues  
Total number of hits satisfying chosen parameters: 583

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%  
Maximum Match 100%  
Listing first 45 summaries

Database : A.Geneseq-101002.\*  
1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	23	100.0	7	21	AAB35994	Sorbitol dehydroge
2	23	100.0	11	9	AAP82045	"Peptide 3" consis
3	23	100.0	11	21	AAV32231	G-CSF agonist mab1
4	23	100.0	13	18	AAW38119	PPPY motif contai
5	23	100.0	13	22	AAE01065	Horse colostrum-as
6	23	100.0	19	20	AAV49212	Seq ID No. 6 of US
7	23	100.0	28	22	ABB39349	Peptide #6855 enco
8	23	100.0	28	22	AAW60026	Human brain marrow
9	23	100.0	28	22	AAW72626	Human bone marrow
10	23	100.0	28	23	ABG42450	Human peptide enco

11	23	100.0	36	22	ABB29961	Peptide #2612 enco
12	23	100.0	36	22	ABB35139	Peptide #2645 enco
13	23	100.0	36	22	ABB20561	Protein #2560 enco
14	23	100.0	36	22	AAW35966	Human Brain expres
15	23	100.0	36	22	AAW68335	Human bone marrow
16	23	100.0	36	22	AAW16152	Peptide #2586 enco
17	23	100.0	36	22	AAW28645	Peptide #2682 enco
18	23	100.0	36	22	AAW03882	Peptide #2564 enco
19	23	100.0	36	23	ABG37890	Human peptide enco
20	23	100.0	42	20	AAV12228	Human 5' EST secre
21	23	100.0	45	22	AAW90285	Human SAA carboxy
22	23	100.0	46	22	AAW90286	Human SAA carboxy
23	23	100.0	47	22	AAW90287	Human SAA carboxy
24	23	100.0	48	22	AAW90288	Human SAA carboxy
25	23	100.0	49	22	AAW90289	Human SAA carboxy
26	23	100.0	50	22	AAW90290	Human SAA carboxy
27	23	100.0	51	22	AAW90291	Human SAA carboxy
28	23	100.0	52	22	AAW90292	Human SAA carboxy
29	23	100.0	53	22	AAW90293	Human SAA carboxy
30	23	100.0	54	22	AAW90294	Human SAA carboxy
31	23	100.0	55	22	AAW90295	Human SAA carboxy
32	23	100.0	56	22	AAW90296	Human SAA carboxy
33	23	100.0	57	22	AAW90297	Human SAA carboxy
34	23	100.0	58	22	AAW90298	Mouse serum amylo1
35	23	100.0	58	22	AAW90298	Human SAA carboxy
36	23	100.0	59	22	AAW90299	Human SAA carboxy
37	23	100.0	60	22	AAW90300	Human SAA carboxy
38	23	100.0	60	22	ABP09824	Human ORFX protein
39	23	100.0	61	22	AAW90301	Human SAA carboxy
40	23	100.0	62	22	AAW90302	Human SAA carboxy
41	23	100.0	63	22	AAW90303	Human SAA carboxy
42	23	100.0	64	22	AAW90304	Human SAA carboxy
43	23	100.0	64	22	AAW90383	Human SAA amino tr
44	23	100.0	65	22	AAW90483	Human SAA internal
45	23	100.0	65	19	AAW62372	Anti-thrombotic pep

## ALIGNMENTS

RESULT 1  
AAB35994  
ID AAB35994 standard; Protein; 7 AA.  
XX  
AC AAB35994;  
XX  
DT 01-MAR-2001 (first entry)  
XX  
DE Sorbitol dehydrogenase subunit 2 internal peptide.  
XX  
KW Sorbitol dehydrogenase; SDH; Gluconobacter oxydans; D-sorbitol;  
KW L-sorbose production; 2-keto-L-gulonic acid.  
XX  
OS Gluconobacter oxydans.  
XX  
PN WO200065066-A1.  
XX  
PD 02-NOV-2000.  
XX  
PF 23-APR-1999; 99WO-IB00736.  
XX  
PR 23-APR-1999; 99WO-IB00736.  
XX  
PA (CHOI/) CHOI E.  
PA (RHEE/) RHEE S.  
PA (LEE/) LEE E.  
PI Choi E, Rhee S, Lee E;  
DR WPI: 2000-687351/67.  
XX  
XX Novel membrane-bound sorbitol dehydrogenase nucleic acid molecule  
PT isolated from Gluconobacter suboxydans useful for the fermentative

PT production of 2-keto-L-gulonic acid and L-sorbose from D-sorbitol -  
 XX  
 PS Example 2; Page 37; 96pp; English.

CC This invention relates to an isolated membrane-bound sorbitol  
 CC dehydrogenase (SDH) from Gluconobacter oxydans. The invention includes  
 CC polynucleotides (AAC83153 - AAC83154) encoding the 3 subunits of SDH  
 CC (AAB35987 - AAB35989). Also included in the invention are two  
 CC polynucleotide sequences AAC83156 and AAC83157 which contain the subunit  
 CC coding regions. The sorbitol dehydrogenase polynucleotide sequences are  
 CC useful for producing L-sorbose from D-sorbitol and for increasing the  
 CC production of 2-keto-L-gulonic acid by transforming a host cell,  
 CC especially Gluconobacter with the DNA and selecting the transformed host  
 CC cell. The present sequence represents the SDH subunit 2 internal  
 CC peptide.

XX  
 SQ Sequence 7 AA;

Query Match 100.0%; Score 23; DB 21; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYDA 4  
 ||||  
 Db 1 NYDA 4

RESULT 2

AAP82045  
 ID AAP82045 standard; peptide; 11 AA.

XX AAP82045;

DT 19-OCT-1990 (first entry)

DE "Peptide 3" consisting of residues 37-47 of amyloid A protein.

XX Amyloid A protein; secondary amyloidosis; anti-amyloid A antibody.

XX synthetic.

XX JP63044895-A.

XX 25-FEB-1988.

XX 13-AUG-1986; 86JP-0189810.

XX 13-AUG-1986; 86JP-0189810.

XX (KYOW ) KYOWA HAKKO KOGYO KK.

XX WPI; 1988-094820/14.

XX Anti-amyloid-A protein monoclonal antibody -

XX used esp for detection of sec amyloidosis

XX Claim 2; Page 649; 6pp; Japanese.

CC An antibody capable of recognising Amyloid A and this peptide  
 CC but which does not react with other specified peptides  
 CC (see AAP82046-9) is useful for detection of secondary amyloidosis.

XX Sequence 11 AA;

Query Match 100.0%; Score 23; DB 9; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 37;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYDA 4  
 ||||  
 Db 5 NYDA 8

RESULT 3  
 ID AAY32231  
 XX AAY32231 standard; Peptide; 11 AA.

XX AAY32231;

DT 15-FEB-2000 (first entry)

DE G-CSF agonist mAb174-74-11 heavy chain variable region CDR3.

XX Granulocyte colony stimulating factor receptor; G-CSF; mouse;

XX monoclonal antibody; agonist; screening; neutropenia; therapy;

XX complementarity determining region; CDR; mAb174-74-11.

XX Mus musculus.

XX WO955735-A1.

XX 04-NOV-1999.

XX 30-APR-1999; 99WO-US09466.

XX 30-APR-1998; 98US-0083575.

XX (TANO-) TANOX INC.

XX Ni B, Sun BNC, Sun CRY;

XX WPI; 2000-052805/04.

XX Treatment of neutropenia by stimulating proliferation of neutrophilic

XX cell lineage progenitors -

XX Claim 14; Page 30; 64pp; English.

XX The present sequence represents complementarity determining region

XX 3 (CDR3) of the heavy chain variable region of murine monoclonal

XX antibody mAb174-74-11. This antibody is an example of an agonist

XX molecule that specifically binds to, or interacts with human

XX granulocyte colony stimulating factor (G-CSF) receptor to stimulate

XX cell proliferation and differentiation, especially by dimerizing

XX the receptor or activating phosphorylation of kinases associated

XX with the receptor. Agonist antibodies can be used to stimulate

XX proliferation of G-CSF-dependent cells, e.g. to differentiate

XX leading to a repopulation of neutrophilic granulocyte lineage

XX cells, especially to treat neutropenia (claimed). They can also

XX be used to detect human G-CSF receptor immunologically.

XX Sequence 11 AA;

Query Match 100.0%; Score 23; DB 21; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 37;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYDA 4  
 ||||  
 Db 5 NYDA 8

RESULT 4  
 ID AAW38119  
 XX AAW38119 standard; Peptide; 13 AA.

XX AAW38119;

DT 23-APR-1998 (first entry)

DE PPPPY motif containing peptide ENAC-beta.

XX Peptide recognition unit; WW domain; cell signalling; growth regulation;

XX cytoskeleton organisation; targeted drug screening; modulator;

XX WW domain interaction; YAP protein; dystrophin;

XX epithelial sodium channel protein.

```

XX OS Synthetic.
XX FH Key
XX FT Region
XX FT 4..8
XX FT /note= "PPPPY motif"
XX PN WO9737223-A1.
XX PD 09-OCT-1997.
XX PF 03-APR-1997; 97WO-US05547.
XX PR 03-APR-1996; 96US-0630916.
XX PA (CYTO-) CYTOGEN CORP.
XX PA (UYNC-) UNIV NORTH CAROLINA.
XX PI Fowlkes DM, Kay BK, Pirozzi G;
XX PI WPI; 1997-503234/46.
XX DR
XX PT Identifying cell signalling and growth regulatory polypeptides by
XX PT reaction with multivalent recognition complex - polypeptides are
XX PT useful in targeted drug selection
XX PS Example 6.3; Fig 24B; 220pp; English.
XX PS
XX CC Peptides AAW38115-28 are derived from epithelial sodium channel alpha
XX CC (ENaC-alpha), beta (ENaC-beta) and gamma (ENaC-gamma) proteins. The
XX CC peptides contain PPPPY-like motifs, which are found in the proline rich
XX CC regions of WBP-1 and WBP-2 proteins. Peptides containing this residue
XX CC bind to the YAP WW domain, but not the WW domain from dystrophin
XX CC or to a panel of SH3 domains. Peptides AAW38115-28 were biotinylated and
XX CC complexed with alkaline streptavidin, and used in a cross affinity
XX CC mapping experiment. They were tested for their ability to bind to the 3
XX CC individual novel WW domains of WWP4 (AAW36797), which were expressed as
XX CC glutathione-S-transferase expression proteins. The present peptide
XX CC binds to some of the WW domains of the WWP4 protein. The WW domain is a
XX CC small functional domain. Its name is derived from the observation that
XX CC two tryptophan residues, one in the amino terminal portion of the WW
XX CC domain and one in the carboxyl terminal portion, are conserved. Most
XX CC proteins containing WW domains have a function involving cell signalling
XX CC and growth regulation or the organisation of the cytoskeleton.
XX CC Polypeptides containing a WW domain are identified by treating a
XX CC multivalent recognition unit complex that has selective binding affinity
XX CC for a WW domain, with many polypeptides and identifying those with
XX CC selective affinity for the complex. Proteins containing WW domains are
XX CC used for targeted drug screening, i.e. to identify potential modulators
XX CC of specific WW domain interactions.
XX SQ Sequence 13 AA:
XX
XX Query Match 100.0%; Score 23; DB 18; Length 13;
XX Best Local Similarity 100.0%; Pred. No. 45;
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 NYDA 4
XX 1111
Db 7 NYDA 10
XX
XX RESULT 5
XX ID AAE01065 standard; peptide; 13 AA.
XX AC AAE01065;
XX XX
XX DE 17-JUL-2001 (first entry)
XX DE Horse colostrum-associated serum amyloid A (SAA) peptide #3.
XX DE Horse; colostrum; serum amyloid A; SAA; antibacterial; antidiarrhoeic;
XX XX

```

```

KM KM apolipoprotein; inflammatory response; mucin 3; MUC3; therapy;
KM KM Escherichia coli infection; traveller's diarrhoea; infant diarrhoea;
KM KM necrotising enterocolitis; NEC; urinary tract infection; dysentery;
XX XX infectious diarrhoea.
XX OS Equus caballus.
XX PN WO200131006-A2.
XX PD 03-MAY-2001.
XX PF 20-OCT-2000; 2000WO-US29065.
XX PR 22-OCT-1999; 99US-0425679.
XX PR 14-JUL-2000; 2000US-0218482.
XX PR 17-JUL-2000; 2000US-0218611.
XX PA (UYNE-) UNIV NEBRASKA.
XX PI McDonald TL, Weber A, Mack DR, Larson MA;
XX PI WPI; 2001-308640/32.
XX DR
XX PT New colostrum-associated serum amyloid A protein induces mucin
XX PT production in epithelial cells and is useful to prevent and treat
XX PT enteric pathogen infection -
XX PS Claim 3; Page 20; 81pp; English.
XX PS
XX CC The patent discloses novel colostrum-associated Serum Amyloid A
XX CC (SAA) from mammals. SAAs are small apolipoproteins that accumulate
XX CC and associate rapidly with high-density lipoprotein 3 (HDL3) during
XX CC the acute phase of inflammatory response. They are characterised
XX CC by the TFLK motif in the N-terminal region which has the ability
XX CC to stimulate mucin 3 (MUC3) production. Colostrum associated SAAs
XX CC are used to prevent or treat infectious diseases associated with
XX CC enteric pathogens (particularly Escherichia coli) such as traveller's
XX CC diarrhoea, infant diarrhoea, necrotising enterocolitis (NEC), urinary
XX CC tract infections, and infectious diarrhoea in herd animals. They are
XX CC used to prevent dysentery and other infectious diseases particularly
XX CC for the military.
XX CC The present sequence is colostrum-associated serum amyloid A (SAA)
XX CC peptide from horse.
XX SQ Sequence 13 AA:
XX
XX Query Match 100.0%; Score 23; DB 22; Length 13;
XX Best Local Similarity 100.0%; Pred. No. 45;
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 NYDA 4
XX 1111
Db 2 NYDA 5
XX
XX RESULT 6
XX ID AAY49212 standard; peptide; 19 AA.
XX AC AAY49212;
XX XX
XX DE 07-FEB-2000 (first entry)
XX DE Seq ID No: 6 of US5977316.
XX DE Monoclonal antibody; MAb; 1A7; GD2; immune response; melanoma;
XX DE neuroblastoma; glioma; soft tissue carcinoma; small cell carcinoma;
XX DE tumor-associated antigen.
XX DE Unidentified.
XX OS US5977316-A.
XX PN
XX XX

```

[illegible]

DR		WPI; 2001-483447/52.		
XX				
PT		Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human fetal liver -		
XX				
PS		Claim 27; SEQ ID NO 31984; 639pp + sequence listing; English.		
XX				
CC		The invention relates to a single exon nucleic acid probe for measuring human gene expression in a sample derived from human foetal liver. The single exon nucleic acid probes may be used for predicting, measuring and displaying gene expression in samples derived from human fetal liver. The present sequence is a peptide encoded by a single exon nucleic acid probe of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.		
CC				
CC				
SC		Sequence     28 AA;		
QY		Query Match		
		Best Local Similarity       100.0%; Score: 23; DB 22; Length 28; Matches      4; Conservative   0; Mismatches   0; Indels     0; Gaps     0;		
Dy		1 NYDA 4		
Db		7 NYDA 10		
RESULT 8				
AAM60026				
ID		AAM60026 standard; Protein; 28 AA.		
AC		AAM60026;		
XX				
DT		05-NOV-2001 (first entry)		
DE				
XX				
KM		Human brain expressed single exon probe encoded protein SEQ ID NO: 32131.		
RW		Human; brain expressed exon; gene expression analysis; probe; microarray; Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.		
OS		Homo sapiens.	.	
PN		WO200157275-A2.		
XX				
PD		09-AUG-2001.		
PE				
PF		30-JAN-2001; 2001WC-USO0667.		
PR				
XX				
PR		04-FEB-2000; 2000US-0180312.		
PR		26-MAY-2000; 2000US-0207456.		
PR		30-JUN-2000; 2000US-0608408.		
PR		03-AUG-2000; 2000US-0632366.		
PR		21-SEP-2000; 2000US-0234687.		
PR		27-SEP-2000; 2000US-0236359.		
PR		04-OCT-2000; 2000GB-0024263.		
PA	(MOLE-) MOLECULAR DYNAMICS INC.			
XX				
PI	Penn SG, Hanzel DK, Chen W, Rank DR;			
DR	WPI; 2001-483446/52.			
XX				
XX				
Example 4; SEQ ID NO: 32131; 650bp + Sequence Listing; English.				
CC		The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human brain. They can be used to measure gene expression in brain cell samples, which may enable the diagnosis and improved treatment of nervous system		

CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
CC epilepsy and cancers. The present sequence is a protein encoded by one of  
XX the probes of the invention.

SO Sequence 28 AA:

Query Match 100.0%; Score 23; DB 22; Length 28;

Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYDA 4  
| | | |  
Db 7 NYDA 10

RESULT 9  
AAM72626

ID AAM72626 standard; Protein; 28 AA.

AC AAM72626;

DT 06-NOV-2001 (first entry)

DE Human bone marrow expressed probe encoded protein SEQ ID NO: 32932.

XX Human: bone marrow expressed exon; gene expression analysis; probe;

KW microarray; cancer; leukemia; lymphoma; myeloma.

XX Homo sapiens.

PN WO200157276-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00668.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human bone marrow -

XX Example 4; SEQ ID NO: 32932; 658bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC bone marrow. They can be used to measure gene expression in bone marrow

CC samples, which may enable the improved diagnosis and treatment of cancers

CC such as lymphoma, leukemia and myeloma. The present sequence is a

CC protein encoded by one of the probes of the invention.

XX Sequence 28 AA:

Query Match 100.0%; Score 23; DB 22; Length 28;

Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYDA 4  
| | | |  
Db 7 NYDA 10

RESULT 10  
ABG42450

ID ABG42450 standard; Peptide; 28 AA.

AC ABG42450;

DT 19-AUG-2002 (first entry)

DE Human peptide encoded by genome-derived single exon probe SEQ ID 32115.

XX Human: single exon probe; asthma; lung cancer; COPD; ILD;

KW chronic obstructive pulmonary disease; interstitial lung disease;

KW familial idiopathic pulmonary fibrosis; neurofibromatosis;

KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;

KW Heremansky-Pudlak syndrome; sarcoidosis; pulmonary haemorrhage;

KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;

KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

KW primary ciliary dyskinesia; pulmonary hypertension;

KW hyaline membrane disease.

XX Homo sapiens.

PN WO200186003-A2.

PD 15-NOV-2001.

PF 30-JAN-2001; 2001WO-US00665.

XX 04-FEB-2000; 2000US-180312P.

PR 26-MAY-2000; 2000US-207456P.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-234687P.

PR 27-SEP-2000; 2000US-236359P.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2002-114183/15.

XX Spatially-addressable set of single exon nucleic acid probes, used to

PT measure gene expression in human lung samples -

XX Claim 27; SEQ ID NO 32115; 634bp; English.

XX The invention relates to a spatially-addressable set of single exon

CC nucleic acid probes for measuring gene expression in a sample derived

CC from human lung comprising single exon nucleic acid probes having one of

CC 12614 nucleic acid sequences mentioned in the specification, or their

CC complements or the 12387 open reading frames derived from the 12614

CC probes; the novel set of probes which hybridize at high stringency to a

CC nucleic acid expressed in the human lung; measuring gene expression in a

CC sample derived from human lung, comprising (a) contacting the array with

CC using human lung derived mRNA and for the study of lung diseases  
CC such as asthma, lung cancer, chronic obstructive pulmonary disease  
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary  
CC fibrosis, neurofibromatosis, tuberosus sclerosis, Gaucher's disease,  
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary  
CC haemangiosarcoma, pulmonary histiocytosis, lymphangioleiomyomatosis,  
CC pulmonary alveolar proteinosis, Karsenger syndrome, fibrocystic  
CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension  
CC and hyaline membrane disease. The present sequence is a peptide/protein  
CC encoded by a single exon probe of the invention.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
CC  
SQ Sequence 28 AA;  
  
Query Match 100.0%; Score 23; DB 23; Length 28;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 NYDA 4  
||||  
Db 7 NYDA 10  
  
RESULT 11  
ABB29961  
ID ABB29961 standard; Peptide; 36 AA.  
XX  
AC ABB29961;  
XX  
DT 01-FEB-2002 (first entry)  
XX  
DE Peptide #2612 encoded by breast cell single exon nucleic acid probe.  
XX  
KM Human; microarray; single exon probe; gene expression; breast;  
KM disease; cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO200157271-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00662.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-496933/54.  
XX  
PT New spatially-addressable set of single exon nucleic acid probes,  
XX useful for measuring gene expression in sample derived from human  
XX breast, comprises number of single exon nucleic acid probes -  
XX  
PS Claim 27; SEQ ID NO 12929; 327pp + sequence listing; English.  
XX  
CC The invention relates to a spatially-addressable set of single exon  
CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human breast and BR 474 cells. The method involves contacting  
CC the probes with a collection of detectably labelled nucleic acids  
CC derived from mRNA of human breast, and then measuring the label  
CC bound to each probe of the microarray. The probes are useful for

CC verifying the expression of regions of genomic DNA predicted to  
CC encode proteins. They are useful for gene discovery, and for  
CC determining predisposition and/or prognosing breast disease. Gene  
CC expression analysis is useful for assessing the toxicity of chemical  
CC agents on cells. The microarray of this invention presents a far greater  
CC diversity of probes for measuring gene expression, with far less bias  
CC than expressed sequence tag microarrays. The method is suitable for  
CC rapid production of functional information from genomic sequence. The  
CC present sequence is a peptide encoded by a single exon nucleic acid  
CC probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
CC  
SQ Sequence 36 AA;  
  
Query Match 100.0%; Score 23; DB 22; Length 36;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 NYDA 4  
||||  
Db 30 NYDA 33  
  
RESULT 12  
ABB35139  
ID ABB35139 standard; Peptide; 36 AA.  
XX  
AC ABB35139;  
XX  
DT 04-FEB-2002 (first entry)  
XX  
DE Peptide #2645 encoded by human foetal liver single exon probe.  
XX  
KM Human; foetal liver; gene expression; single exon nucleic acid probe.  
XX  
OS Homo sapiens.  
XX  
PN WO200157277-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00669.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-483447/52.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
XX analyzing gene expression in human foetal liver -  
XX  
PS Claim 27; SEQ ID NO 27774; 639pp + sequence listing; English.  
XX  
CC The invention relates to a single exon nucleic acid probe for  
CC measuring human gene expression in a sample derived from human foetal  
CC liver. The single exon nucleic acid probes may be used for predicting,  
CC measuring and displaying gene expression in samples derived from human  
CC foetal liver. The present sequence is a peptide encoded by a single exon  
CC nucleic acid probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.



```

XX      SQ      Sequence      36 AA;
Query Match      100.0%; Score 23; DB 22; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY      1 NYDA 4
      ||||
Db      30 NYDA 33

RESULT 13
ABB20561
ID      ABB20561 standard; Protein: 36 AA.
AC      ABB20561;
XX
XX      DT      23-JAN-2002 (first entry)
XX
DE      Protein #2560 encoded by probe for measuring heart cell gene expression.
XX
XX      KW      Human: gene expression; heart; microarray; vascular system;
KW      cardiovascular disease; hypertension; cardiac arrhythmia;
XX      congenital heart disease.
XX
OS      Homo sapiens.
XX
XX      PN      WO200157274-A2.
XX
XX      PD      09-AUG-2001.
XX
XX      PF      30-JAN-2001; 2001WO-US00666.
XX
XX      PR      04-FEB-2000; 2000US-0180312.
XX      PR      26-MAY-2000; 2000US-0207456.
XX      PR      30-JUN-2000; 2000US-0608408.
XX      PR      03-AUG-2000; 2000US-0632366.
XX      PR      21-SEP-2000; 2000US-0234687.
XX      PR      27-SEP-2000; 2000US-0236359.
XX      PR      04-OCT-2000; 2000GB-0024263.
XX
XX      PA      (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX      PI      Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX      DR      WPI; 2001-488899/53.
XX
XX      PT      Single exon nucleic acid probes for analyzing gene expression in human
XX      hearts -
XX
XX      PS      Claim 15; SEQ ID NO 22331; 530pp; English.
XX
XX      CC      The present invention relates to single exon nucleic acid probes for
XX      measuring human gene expression in a sample derived from human heart (see
XX      ABA21535-ABA41305). The present sequence is a protein encoded by one such
XX      probe. The probes may be used for predicting, measuring and displaying
XX      gene expression in samples derived from the human heart via microarrays.
XX      By measuring gene expression, the probes are useful for predicting,
XX      diagnosing, grading, staging, monitoring and prognosing diseases of the
XX      human heart and vascular system e.g. cardiovascular disease,
XX      hypertension, cardiac arrhythmias and congenital heart disease.
XX      Note: The sequence data for this patent did not form part of the printed
XX      specification, but was obtained in electronic format directly from WIPO
XX      at ftp.wipo.int/pub/published_pct_sequences.
XX
XX      SQ      Sequence      36 AA;
Query Match      100.0%; Score 23; DB 22; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY      1 NYDA 4

```

```

Db      30 NYDA 33
      ||||
RESULT 14
AAM55966
ID      AAM55966 standard; Protein: 36 AA.
AC      AAM55966;
XX
XX      DT      05-NOV-2001 (first entry)
XX
XX      DE      Human brain expressed single exon probe encoded protein SEQ ID NO: 28071.
XX      KW      Human: brain expressed exon; gene expression analysis; probe;
XX      microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX      epilepsy; cancer.
XX
XX      OS      Homo sapiens.
XX
XX      PN      WO200157275-A2.
XX
XX      PD      09-AUG-2001.
XX
XX      PF      30-JAN-2001; 2001WO-US00667.
XX
XX      PR      04-FEB-2000; 2000US-0180312.
XX      PR      26-MAY-2000; 2000US-0207456.
XX      PR      30-JUN-2000; 2000US-0608408.
XX      PR      03-AUG-2000; 2000US-0632366.
XX      PR      21-SEP-2000; 2000US-0234687.
XX      PR      27-SEP-2000; 2000US-0236359.
XX      PR      04-OCT-2000; 2000GB-0024263.
XX
XX      PA      (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX      PI      Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX      DR      WPI; 2001-483446/52.
XX
XX      PT      Single exon nucleic acid probes for analyzing gene expression in human
XX      brains -
XX
XX      PS      Example 4; SEQ ID NO: 28071; 650pp + Sequence Listing; English.
XX
XX      CC      The present invention provides a number of single exon nucleic acid
XX      probes which are derived from genomic sequences expressed in the human
XX      brain. They can be used to measure gene expression in brain cell samples,
XX      which may enable the diagnosis and improved treatment of nervous system,
XX      diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX      epilepsy and cancers. The present sequence is a protein encoded by one of
XX      the probes of the invention.
XX
XX      SQ      Sequence      36 AA;
Query Match      100.0%; Score 23; DB 22; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY      1 NYDA 4
      ||||
Db      30 NYDA 33

RESULT 15
AAM68335
ID      AAM68335 standard; Protein: 36 AA.
AC      AAM68335;
XX
XX      DT      06-NOV-2001 (first entry)
XX
XX      DE      Human bone marrow expressed probe encoded protein SEQ ID NO: 28641.

```

XX Human; bone marrow expressed exon; gene expression analysis; probe;  
KW microarray; cancer; leukemia; lymphoma; myeloma.  
XX Homo sapiens.  
XX WO200157276-A2.  
XX  
XX 09-AUG-2001.  
XX  
XX 30-JAN-2001; 2001WO-US00668.  
XX  
XX 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-488900/53.  
XX  
XX Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human bone marrow -  
XX  
XX Example 4; SEQ ID NO: 28641; 658pp + Sequence Listing; English.  
XX  
XX The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC bone marrow. They can be used to measure gene expression in bone marrow  
CC samples, which may enable the improved diagnosis and treatment of cancers  
CC such as lymphoma, leukemia and myeloma. The present sequence is a  
CC protein encoded by one of the probes of the invention.  
XX  
SQ Sequence 36 AA;  
  
Query Match 100.0%; Score 23; DB 22; Length 36;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 NYDA 4  
| | | |  
Db 30 NYDA 33  
  
RESULT 16  
AAM16152  
ID AAM16152 standard; Protein; 36 AA.  
XX  
XX AAM16152;  
XX  
XX 12-OCT-2001 (first entry)  
XX  
DE Peptide #2586 encoded by probe for measuring cervical gene expression.  
XX  
XX Probe; human; microarray; gene expression; cervical epithelial cell;  
KW cervical cancer.  
XX  
XX Homo sapiens.  
XX  
XX WO200157278-A2.  
XX  
XX 09-AUG-2001.  
XX  
XX 30-JAN-2001; 2001WO-US00670.  
XX  
XX 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-488900/53.  
XX

PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-488901/53.  
XX  
XX Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human cervical epithelial cells -  
XX  
XX Claim 27; SEQ ID NO 20978; 487pp; English.  
XX  
XX The present invention relates to human single exon nucleic acid probes  
CC (SNP: see AAI10068-AA128459). The present sequence is a peptide encoded  
CC by one such probe. The SNPs are derived from human HeLa cells. The SNPs  
CC can be used to produce a single exon microarray, which can be used for  
CC measuring human gene expression in a sample derived from human cervical  
CC epithelial cells. By measuring gene expression, the probes are therefore  
CC useful in grading and/or staging of diseases of the cervix, notably  
CC cervical cancer.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 36 AA;  
  
Query Match 100.0%; Score 23; DB 22; Length 36;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 NYDA 4  
| | | |  
Db 30 NYDA 33  
  
RESULT 17  
AAM28645  
ID AAM28645 standard; Protein; 36 AA.  
XX  
XX AAM28645;  
XX  
XX 17-OCT-2001 (first entry)  
XX  
DE Peptide #2682 encoded by probe for measuring placental gene expression.  
XX  
XX Probe; microarray; human; placenta; antenatal diagnosis;  
KW genetic disorder.  
XX  
XX Homo sapiens.  
XX  
XX WO200157272-A2.  
XX  
XX 09-AUG-2001.  
XX  
XX 30-JAN-2001; 2001WO-US00663.  
XX  
XX 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-48897/53.  
XX

XX Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human placenta -  
XX  
XX Claim 27; SEQ ID No 28914; 654pp; English.  
XX  
CC The present invention relates to single exon nucleic acid probes (SENP;  
CC see A11315-A15746). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for producing a microarray for  
CC predicting, measuring and displaying gene expression in samples derived  
CC from human placenta. The probes are useful for antenatal diagnosis of  
CC human genetic disorders.  
XX  
SQ Sequence 36 AA;  
XX  
Query Match 100.0%; Score 23; DB 22; Length 36;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NYDA 4  
Db 30 NYDA 33  
XX  
RESULT 18  
AAM03882  
ID AAM03882 standard; Protein: 36 AA.  
XX  
AC AAM03882;  
XX  
DT 09-OCT-2001 (first entry)  
XX  
DE Peptide #2554 encoded by probe for measuring breast gene expression.  
XX  
KW Probe: human; breast disease; breast cancer; development disorder;  
XX Inflammatory disease; proliferative breast disease; non-carcinoma tumour.  
XX  
OS Homo sapiens.  
XX  
PN WO200157270-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 29-JAN-2001; 2001WO-US00661.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX WPI; 2001-476286/51.  
XX  
DR Novel single exon nucleic acid probe used to measuring gene expression  
XX in a human breast -  
XX  
XX Claim 27; SEQ ID No 12622; 322pp; English.  
XX  
CC The present invention relates to novel single exon nucleic acid probes  
CC (see A1100010-A110067). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for measuring human gene expression in  
CC a human breast sample, where the probe hybridizes at high stringency to a  
CC nucleic acid expressed in the human breast. The probes are useful for  
CC predicting, diagnosing, grading, staging, monitoring and prognosing  
CC diseases of the human breast, particularly those diseases with polygenic  
CC etiology. The diseases include: breast cancer, disorders of development,  
CC inflammatory diseases of the breast, fibrocytic changes, proliferative

CC breast disease and non-carcinoma tumours.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 36 AA;  
XX  
Query Match 100.0%; Score 23; DB 22; Length 36;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NYDA 4  
Db 30 NYDA 33  
XX  
RESULT 19  
ABG37890  
ID ABG37890 standard; Peptide: 36 AA.  
XX  
AC ABG37890;  
XX  
DT 19-AUG-2002 (first entry)  
XX  
DE Human peptide encoded by genome-derived single exon probe SEQ ID 27555.  
XX  
KW Human; single exon probe; asthma; lung cancer; COPD; ILD;  
KW chronic obstructive pulmonary disease; interstitial lung disease;  
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;  
KW tuberculous sclerosis; Gaucher's disease; Niemann-Pick disease;  
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karageners syndrome;  
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
KW primary ciliary dyskinesia; pulmonary hypertension;  
XX  
XX hyaline membrane disease.  
XX  
OS Homo sapiens.  
XX  
PN WO200186003-A2.  
XX  
PD 15-NOV-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00665.  
XX  
PR 04-FEB-2000; 2000US-180312P.  
PR 26-MAY-2000; 2000US-207456P.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-234687P.  
PR 27-SEP-2000; 2000US-236359P.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX WPI; 2002-114183/15.  
XX  
DR Spatially-addressable set of single exon nucleic acid probes, used to  
XX measure gene expression in human lung samples -  
XX  
XX Claim 27; SEQ ID No 27555; 634pp; English.  
XX  
CC The invention relates to a spatially-addressable set of single exon  
CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human lung comprising single exon nucleic acid probes having one of  
CC 12614 nucleic acid sequences mentioned in the specification, or their  
CC complements or the 12387 open reading frames derived from the 12614  
CC probes. Also included are a microarray comprising the novel set of  
CC probes; the novel set of probes which hybridise at high stringency to a  
CC nucleic acid expressed in the human lung; measuring gene expression in a  
CC sample derived from human lung, comprising (a) contacting the array with  
CC a collection of detectably labeled nucleic acids derived from human lung

CC mRNA, and (b) measuring the label detectably bound to each probe of  
 CC the array; identifying exons in a eukaryotic genome, comprising  
 CC (a) algorithmically predicting at least one exon from genomic sequences  
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably  
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
 CC having a fragment identical to the predicted exon, the probe is included  
 CC in the above mentioned microarray; assigning exons to a single gene,  
 CC comprising (a) identifying exons from genomic sequence by the method  
 CC above and (b) measuring the expression of each of the exons in several  
 CC tissues and/or cell types using hybridisation to a single exon  
 CC microarrays having a probe with the exon, where a common pattern of  
 CC expression of the exons in the tissues and/or cell types indicates that  
 CC the exons should be assigned to a single gene; a peptide comprising one  
 CC of 1201 sequences, mentioned in the specification, or encoded by the  
 CC probes/open reading frames (ORF). The probes are used for gene  
 CC expression analysis, and for identifying exons in a gene, particularly  
 CC using human lung derived mRNA and for the study of lung diseases  
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease  
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary  
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,  
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary  
 CC hemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,  
 CC pulmonary alveolar proteinosis, Karagenet syndrome, fibrocystic  
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension  
 CC and hyaline membrane disease. The present sequence is a peptide/protein  
 CC encoded by a single exon probe of the invention.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 36 AA:

Query Match 100.0%; Score 23; DB 23; Length 36;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYDA 4  
 ||||  
 Db 30 NYDA 33

RESULT 20

AA12228  
 ID AAY12228 standard; Protein; 42 AA.

AC AAY12228;

DT 18-JUN-1999 (first entry)

DE Human 5' EST secreted protein SEQ ID NO: 541.

XX Human: secreted protein; EST: expressed sequence tag; diagnosis;  
 XX forensic: gene therapy; chromosome mapping; signal peptide;  
 XX upstream regulatory sequence; cytokine activity; cell proliferation;  
 XX differentiation; haematopoiesis regulation; tissue growth regulation;  
 XX reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;  
 XX thrombolytic; anti-inflammatory; tumour inhibition.

OS Homo sapiens.

PN WO9906554-A2.

PD 11-FEB-1999.

PF 31-JUL-1998; 98WO-IB01238.

PR 01-AUG-1997; 97US-0905134.

PI (GEST ) GENSET.

PA Duclert A, Dumas Milne Edwards J, Lacroix B;  
 XX PT

XX WPI: 1999-153784/13.  
 DR N-PSDB: AAY41061.

PT New nucleic acids encoding human secreted proteins - obtained from  
 PT cDNA libraries prepared from kidney, fetal kidney, dystrophic  
 PT muscle, muscle and heart tissue

PS Claim 34; Page 602; 622pp; English.

CC AAX40826 to AAX41093 represent 5' expressed sequence tags (ESTs) for  
 CC human secreted proteins, and encode the proteins given in AAY01602 and  
 CC AAY11994 to AAY12260, respectively. The proteins given represent the  
 CC signal peptide and an N-terminal fragment of a secreted protein. The  
 CC nucleic acid sequences can be used for producing secreted human gene  
 CC products. They can also be used to develop products for diagnosis and  
 CC therapy. The proteins obtained may have cytokine activity, cell  
 CC proliferation/differentiation activity, haematopoiesis regulating  
 CC activity, tissue growth regulating activity, reproductive hormone  
 CC regulating activity, chemotactic/chemokinetic activity, haemostatic and  
 CC thrombolytic activity, receptor/ligand activity, anti-inflammatory  
 CC activity, tumour inhibition activity or other activities. The products  
 CC can be used in forensic, gene therapy and chromosome mapping procedures.  
 CC The sequences can also be used for obtaining corresponding promoter  
 CC sequences. The nucleic acids encoding the signal peptide can be used  
 CC for directing extracellular secretion of a polypeptide or the insertion  
 CC of a polypeptide into a membrane, or importing a polypeptide into  
 CC a cell.

Sequence 42 AA:

Query Match 100.0%; Score 23; DB 20; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYDA 4  
 ||||  
 Db 31 NYDA 34

RESULT 21

AAB90285  
 ID AAB90285 standard; Protein; 45 AA.

AC AAB90285;

DT 01-JUN-2001 (first entry)

DE Human SAA carboxy truncation, SEQ ID NO: 43.

XX Human: serum amyloid A; SAA; human formyl peptide receptor variant;  
 XX FPR1; immunomodulatory; anti-inflammatory; antimicrobial; antiarthritic;  
 XX antiatherosclerotic; immunosuppressive; SAA/FPR1 complex; cancer;  
 XX immune system disorder; amyloidosis; inflammation; infection;  
 XX organ rejection; arthritis; atherosclerosis.

OS Homo sapiens.

PN WO200121188-A1.

PD 29-MAR-2001.

PF 22-SEP-1999; 99WO-US21770.

PR 22-SEP-1999; 99WO-US21770.

PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

PI Wang J, Oppenheim JJ, Su S, Gong W, Gao J, Murphy PW;

DR WPI: 2001-211457/21.

XX New serum amyloid A and formyl peptide receptor variant complex and its  
 PT

PT modulators, useful for treating immune system disorders, amyloidosis,  
 PT inflammation, infection, organ rejection, arthritis, atherosclerosis,  
 PT and neoplasia -  
 PS Disclosure: Page 15; 141pp; English.  
 CC The present sequence is a fragment of human serum amyloid A (SAA). The  
 CC invention relates to the discovery that SAA is a ligand for FPR1 (human  
 CC formyl peptide receptor variant). A complex has been isolated that  
 CC comprises a peptide agent with a sequence corresponding to SAA, or its  
 CC conservative variant or functional fragment, bound to FPR1. Modulators  
 CC of the SAA/FPR1 complex are useful for treating immune system  
 CC disorders, amyloidosis, inflammation, infection, organ rejection,  
 CC arthritis, atherosclerosis and neoplasia. Fragments of SAA were tested  
 CC for their ability to modulate assembly of the SAA/FPR1 complex.  
 CC  
 SQ Sequence 45 AA;  
 Query Match 100.0%; Score 23; DB 22; Length 45;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NYDA 4  
 DB 42 NYDA 45  
 RESULT 22  
 AAB90286  
 ID AAB90286 standard; Protein; 46 AA.  
 AC AAB90286;  
 XX  
 DT 01-JUN-2001 (first entry)  
 DE Human SAA carboxy truncation, SEQ ID NO: 44.  
 XX  
 DE Human SAA carboxy truncation, SEQ ID NO: 44.  
 XX  
 KW Human; serum amyloid A; SAA; human formyl peptide receptor variant;  
 KW FPR1; immunomodulatory; antiinflammatory; antimicrobial; antiarthritic;  
 KW antiatherosclerotic; immunosuppressive; SAA/FPR1 complex; cancer;  
 KW immune system disorder; amyloidosis; inflammation; infection;  
 KW organ rejection; arthritis; atherosclerosis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200121188-A1.  
 XX  
 PD 29-MAR-2001.  
 XX  
 PF 22-SEP-1999; 99WO-US21770.  
 XX  
 PR 22-SEP-1999; 99WO-US21770.  
 XX  
 PS (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PI Wang J, Oppenheim JJ, Su S, Gong W, Gao J, Murphy PM.  
 XX  
 DR WPI; 2001-211457/21.  
 XX  
 PT New serum amyloid A and formyl peptide receptor variant complex and its  
 PT modulators, useful for treating immune system disorders, amyloidosis,  
 PT inflammation, infection, organ rejection, arthritis, atherosclerosis  
 PT and neoplasia -  
 PS Disclosure: Page 15; 141pp; English.  
 XX  
 PS The present sequence is a fragment of human serum amyloid A (SAA). The  
 CC invention relates to the discovery that SAA is a ligand for FPR1 (human  
 CC formyl peptide receptor variant). A complex has been isolated that  
 CC comprises a peptide agent with a sequence corresponding to SAA, or its  
 CC conservative variant or functional fragment, bound to FPR1. Modulators  
 CC of the SAA/FPR1 complex are useful for treating immune system  
 CC disorders, amyloidosis, inflammation, infection, organ rejection,  
 CC arthritis, atherosclerosis and neoplasia. Fragments of SAA were tested  
 CC for their ability to modulate assembly of the SAA/FPR1 complex.

CC arthritis, atherosclerosis and neoplasia. Fragments of SAA were tested  
 CC for their ability to modulate assembly of the SAA/FPR1 complex.  
 CC  
 SQ Sequence 46 AA;  
 Query Match 100.0%; Score 23; DB 22; Length 46;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NYDA 4  
 DB 42 NYDA 45  
 RESULT 23  
 AAB90287  
 ID AAB90287 standard; Protein; 47 AA.  
 XX  
 AC AAB90287;  
 XX  
 DT 01-JUN-2001 (first entry)  
 DE Human SAA carboxy truncation, SEQ ID NO: 45.  
 XX  
 DE Human; serum amyloid A; SAA; human formyl peptide receptor variant;  
 KW FPR1; immunomodulatory; antiinflammatory; antimicrobial; antiarthritic;  
 KW antiatherosclerotic; immunosuppressive; SAA/FPR1 complex; cancer;  
 KW immune system disorder; amyloidosis; inflammation; infection;  
 KW organ rejection; arthritis; atherosclerosis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200121188-A1.  
 XX  
 PD 29-MAR-2001.  
 XX  
 PF 22-SEP-1999; 99WO-US21770.  
 XX  
 PR 22-SEP-1999; 99WO-US21770.  
 XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PI Wang J, Oppenheim JJ, Su S, Gong W, Gao J, Murphy PM.  
 XX  
 DR WPI; 2001-211457/21.  
 XX  
 PT New serum amyloid A and formyl peptide receptor variant complex and its  
 PT modulators, useful for treating immune system disorders, amyloidosis,  
 PT inflammation, infection, organ rejection, arthritis, atherosclerosis  
 PT and neoplasia -  
 PS Disclosure: Page 15; 141pp; English.  
 XX  
 PS The present sequence is a fragment of human serum amyloid A (SAA). The  
 CC invention relates to the discovery that SAA is a ligand for FPR1 (human  
 CC formyl peptide receptor variant). A complex has been isolated that  
 CC comprises a peptide agent with a sequence corresponding to SAA, or its  
 CC conservative variant or functional fragment, bound to FPR1. Modulators  
 CC of the SAA/FPR1 complex are useful for treating immune system  
 CC disorders, amyloidosis, inflammation, infection, organ rejection,  
 CC arthritis, atherosclerosis and neoplasia. Fragments of SAA were tested  
 CC for their ability to modulate assembly of the SAA/FPR1 complex.  
 CC  
 SQ Sequence 47 AA;  
 Query Match 100.0%; Score 23; DB 22; Length 47;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NYDA 4  
 DB 42 NYDA 45

```

RESULT 24
AAB90288
ID AAB90288 standard; Protein; 48 AA.
AC AAB90288;
XX
XX
DT 01-JUN-2001 (first entry)
XX
XX
DE Human SAA carboxy truncation, SEQ ID NO: 46.
XX
XX
KW Human; serum amyloid A; SAA; human formyl peptide receptor variant;
KW FPR1; immunomodulatory; antiinflammatory; antimicrobial; antiarthritic;
KW antiatherosclerotic; immunosuppressive; SAA/FPR1 complex; cancer;
KW immune system disorder; amyloidosis; inflammation; infection;
KW organ rejection; arthritis; atherosclerosis.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200121188-A1.
XX
XX
PD 29-MAR-2001.
XX
XX
PE 22-SEP-1999; 99WO-US21770.
XX
XX
PR 22-SEP-1999; 99WO-US21770.
XX
XX
PS (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX
PI Wang J, Oppenheim J, Su S, Gong W, Gao J, Murphy PM;
XX
XX
DR WPI; 2001-211457/21.
XX
XX
PT New serum amyloid A and formyl peptide receptor variant complex and its
PT modulators, useful for treating immune system disorders, amyloidosis,
PT inflammation, infection, organ rejection, arthritis, atherosclerosis
PT and neoplasia -
XX
XX
PS Disclosure; Page 15; 141pp; English.
XX
XX
CC The present sequence is a fragment of human serum amyloid A (SAA). The
CC invention relates to the discovery that SAA is a ligand for FPR1 (human
CC formyl peptide receptor variant). A complex has been isolated that
CC comprises a peptide agent with a sequence corresponding to SAA, or its
CC conservative variant or functional fragment, bound to FPR1. Modulators
CC of the SAA/FPR1 complex are useful for treating immune system
CC disorders, amyloidosis, inflammation, infection, organ rejection,
CC arthritis, atherosclerosis and neoplasia. Fragments of SAA were tested
CC for their ability to modulate assembly of the SAA/FPR1 complex.
XX
XX
SQ Sequence 48 AA;
XX
XX
Query Match 100.0%; Score 23; DB 22; Length 48;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 NYDA 4
DB 42 NYDA 45
XX
XX
DE 01-JUN-2001 (first entry)
XX
XX
DE Human SAA carboxy truncation, SEQ ID NO: 47.
XX
XX
KW Human; serum amyloid A; SAA; human formyl peptide receptor variant;
KW FPR1; immunomodulatory; antiinflammatory; antimicrobial; antiarthritic;
KW organ rejection; arthritis; atherosclerosis.

```

```

KW antiatherosclerotic; immunosuppressive; SAA/FPR1 complex; cancer;
KW immune system disorder; amyloidosis; inflammation; infection;
KW organ rejection; arthritis; atherosclerosis.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200121188-A1.
XX
XX
PD 29-MAR-2001.
XX
XX
PE 22-SEP-1999; 99WO-US21770.
XX
XX
PR 22-SEP-1999; 99WO-US21770.
XX
XX
PS (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX
PI Wang J, Oppenheim J, Su S, Gong W, Gao J, Murphy PM;
XX
XX
DR WPI; 2001-211457/21.
XX
XX
PT New serum amyloid A and formyl peptide receptor variant complex and its
PT modulators, useful for treating immune system disorders, amyloidosis,
PT inflammation, infection, organ rejection, arthritis, atherosclerosis
PT and neoplasia -
XX
XX
PS Disclosure; Page 15; 141pp; English.
XX
XX
CC The present sequence is a fragment of human serum amyloid A (SAA). The
CC invention relates to the discovery that SAA is a ligand for FPR1 (human
CC formyl peptide receptor variant). A complex has been isolated that
CC comprises a peptide agent with a sequence corresponding to SAA, or its
CC conservative variant or functional fragment, bound to FPR1. Modulators
CC of the SAA/FPR1 complex are useful for treating immune system
CC disorders, amyloidosis, inflammation, infection, organ rejection,
CC arthritis, atherosclerosis and neoplasia. Fragments of SAA were tested
CC for their ability to modulate assembly of the SAA/FPR1 complex.
XX
XX
SQ Sequence 49 AA;
XX
XX
Query Match 100.0%; Score 23; DB 22; Length 49;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 NYDA 4
DB 42 NYDA 45
XX
XX
DE 01-JUN-2001 (first entry)
XX
XX
DE Human SAA carboxy truncation, SEQ ID NO: 48.
XX
XX
KW Human; serum amyloid A; SAA; human formyl peptide receptor variant;
KW FPR1; immunomodulatory; antiinflammatory; antimicrobial; antiarthritic;
KW antiatherosclerotic; immunosuppressive; SAA/FPR1 complex; cancer;
KW immune system disorder; amyloidosis; inflammation; infection;
KW organ rejection; arthritis; atherosclerosis.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200121188-A1.
XX
XX
PD 29-MAR-2001.
XX
XX
PE 22-SEP-1999; 99WO-US21770.
XX
XX
PR 22-SEP-1999; 99WO-US21770.

```

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
PA Wang J, Oppenheim JJ, Su S, Gong W, Gao J, Murphy PM;  
XX WPI; 2001-211457/21.  
DR  
XX New serum amyloid A and formyl peptide receptor variant complex and its  
PT modulators, useful for treating immune system disorders, amyloidosis,  
PT inflammation, infection, organ rejection, arthritis, atherosclerosis  
PT and neoplasia -  
XX  
PS Disclosure; Page 15; 141pp; English.  
XX  
XX The present sequence is a fragment of human serum amyloid A (SAA). The  
CC invention relates to the discovery that SAA is a ligand for FPR1 (human  
CC formyl peptide receptor variant). A complex has been isolated that  
CC comprises a peptide agent with a sequence corresponding to SAA, or its  
CC conservative variant or functional fragment, bound to FPR1. Modulators  
CC of the SAA/FPR1 complex are useful for treating immune system  
CC disorders, amyloidosis, inflammation, infection, organ rejection,  
CC arthritis, atherosclerosis and neoplasia. Fragments of SAA were tested  
CC for their ability to modulate assembly of the SAA/FPR1 complex.  
CC  
XX  
SQ Sequence 50 AA;  
OY  
Query Match 100.0%; Score 23; DB 22; Length 50;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 NYDA 4  
1111  
DB 42 NYDA 45  
RESULT 27  
AAB90291  
ID AAB90291 standard; Protein; 51 AA.  
XX  
XX AAB90291;  
AC  
XX  
DT 01-JUN-2001 (first entry)  
DE  
XX Human SAA carboxy truncation, SEQ ID NO: 49.  
DE  
XX  
XX Human; serum amyloid A; SAA; human formyl peptide receptor variant;  
KW FPR1; immunomodulatory; antiinflammatory; antimicrobial; antiarthritic;  
KW antiatherosclerotic; immunosuppressive; SAA/FPR1 complex; cancer;  
KW immune system disorder; amyloidosis; inflammation; infection;  
KW organ rejection; arthritis; atherosclerosis.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200121188-A1.  
PN  
XX  
XX 29-MAR-2001.  
PD  
XX  
XX 22-SEP-1999; 99WO-US21770.  
PE  
XX  
XX 22-SEP-1999; 99WO-US21770.  
PR  
XX  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
PA Wang J, Oppenheim JJ, Su S, Gong W, Gao J, Murphy PM;  
XX WPI; 2001-211457/21.  
DR  
XX New serum amyloid A and formyl peptide receptor variant complex and its  
PT modulators, useful for treating immune system disorders, amyloidosis,  
PT inflammation, infection, organ rejection, arthritis, atherosclerosis  
PT and neoplasia -  
XX  
PS Disclosure; Page 15; 141pp; English.

XX  
XX The present sequence is a fragment of human serum amyloid A (SAA). The  
CC invention relates to the discovery that SAA is a ligand for FPR1 (human  
CC formyl peptide receptor variant). A complex has been isolated that  
CC comprises a peptide agent with a sequence corresponding to SAA, or its  
CC conservative variant or functional fragment, bound to FPR1. Modulators  
CC of the SAA/FPR1 complex are useful for treating immune system  
CC disorders, amyloidosis, inflammation, infection, organ rejection,  
CC arthritis, atherosclerosis and neoplasia. Fragments of SAA were tested  
CC for their ability to modulate assembly of the SAA/FPR1 complex.  
CC  
XX  
SQ Sequence 51 AA;  
OY  
Query Match 100.0%; Score 23; DB 22; Length 51;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 NYDA 4  
1111  
DB 42 NYDA 45  
RESULT 28  
AAB90292  
ID AAB90292 standard; Protein; 52 AA.  
XX  
XX AAB90292;  
AC  
XX  
DT 01-JUN-2001 (first entry)  
DE  
XX Human SAA carboxy truncation, SEQ ID NO: 50.  
DE  
XX  
XX Human; serum amyloid A; SAA; human formyl peptide receptor variant;  
KW FPR1; immunomodulatory; antiinflammatory; antimicrobial; antiarthritic;  
KW antiatherosclerotic; immunosuppressive; SAA/FPR1 complex; cancer;  
KW immune system disorder; amyloidosis; inflammation; infection;  
KW organ rejection; arthritis; atherosclerosis.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200121188-A1.  
PN  
XX  
XX 29-MAR-2001.  
PD  
XX  
XX 22-SEP-1999; 99WO-US21770.  
PE  
XX  
XX 22-SEP-1999; 99WO-US21770.  
PR  
XX  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
PA Wang J, Oppenheim JJ, Su S, Gong W, Gao J, Murphy PM;  
XX WPI; 2001-211457/21.  
DR  
XX  
XX New serum amyloid A and formyl peptide receptor variant complex and its  
PT modulators, useful for treating immune system disorders, amyloidosis,  
PT inflammation, infection, organ rejection, arthritis, atherosclerosis  
PT and neoplasia -  
XX  
XX  
PS Disclosure; Page 15; 141pp; English.  
XX  
XX The present sequence is a fragment of human serum amyloid A (SAA). The  
CC invention relates to the discovery that SAA is a ligand for FPR1 (human  
CC formyl peptide receptor variant). A complex has been isolated that  
CC comprises a peptide agent with a sequence corresponding to SAA, or its  
CC conservative variant or functional fragment, bound to FPR1. Modulators  
CC of the SAA/FPR1 complex are useful for treating immune system  
CC disorders, amyloidosis, inflammation, infection, organ rejection,  
CC arthritis, atherosclerosis and neoplasia. Fragments of SAA were tested  
CC for their ability to modulate assembly of the SAA/FPR1 complex.  
CC  
XX  
SQ Sequence 52 AA;

Query Match 100.0%; Score 23; DB 22; Length 52;  
 Best Local Similarity 100.0%; Pred. No. 2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYDA 4  
 ||||  
 Db 42 NYDA 45

## RESULT 29

AAB90293  
 ID AAB90293 standard; Protein; 53 AA.

AC AAB90293;

DT 01-JUN-2001 (first entry)

DE Human SAA carboxy truncation, SEQ ID NO: 51.

KW Human; serum amyloid A; SAA; human formyl peptide receptor variant;

KW FPR1; immunomodulatory; antiinflammatory; antimicrobial; antiarthritic;

KW antiatherosclerotic; immunosuppressive; SAA/FPR1 complex; cancer;

KW immune system disorder; amyloidosis; inflammation; infection;

KW organ rejection; arthritis; atherosclerosis.

OS Homo sapiens.

PN WO200121188-A1.

PD 29-MAR-2001.

PF 22-SEP-1999; 99WO-US21770.

PR 22-SEP-1999; 99WO-US21770.

PI (USSH ) US DEPT HEALTH & HUMAN SERVICES.

PI Wang J, Oppenheim JJ, Su S, Gong W, Gao J, Murphy PM;

PI WPI; 2001-211457/21.

PT New serum amyloid A and formyl peptide receptor variant complex and its

PT modulators, useful for treating immune system disorders, amyloidosis,

PT inflammation, infection, organ rejection, arthritis, atherosclerosis

PT and neoplasia -

PS Disclosure; Page 15; 141pp; English.

CC The present sequence is a fragment of human serum amyloid A (SAA). The

CC invention relates to the discovery that SAA is a ligand for FPR1 (human

CC formyl peptide receptor variant). A complex has been isolated that

CC comprises a peptide agent with a sequence corresponding to SAA, or its

CC conservative variant or functional fragment, bound to FPR1. Modulators

CC of the SAA/FPR1 complex are useful for treating immune system

CC disorders, amyloidosis, inflammation, infection, organ rejection,

CC arthritis, atherosclerosis and neoplasia. Fragments of SAA were tested

CC for their ability to modulate assembly of the SAA/FPR1 complex.

SQ Sequence 53 AA;

Query Match 100.0%; Score 23; DB 22; Length 53;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYDA 4  
 ||||  
 Db 42 NYDA 45

## RESULT 30

AAB90294  
 ID AAB90294 standard; Protein; 54 AA.

AC AAB90294;

DT 01-JUN-2001 (first entry)

DE Human SAA carboxy truncation, SEQ ID NO: 52.

KW Human; serum amyloid A; SAA; human formyl peptide receptor variant;

KW FPR1; immunomodulatory; antiinflammatory; antimicrobial; antiarthritic;

KW antiatherosclerotic; immunosuppressive; SAA/FPR1 complex; cancer;

KW immune system disorder; amyloidosis; inflammation; infection;

KW organ rejection; arthritis; atherosclerosis.

OS Homo sapiens.

PN WO200121188-A1.

PD 29-MAR-2001.

PF 22-SEP-1999; 99WO-US21770.

PR 22-SEP-1999; 99WO-US21770.

PI (USSH ) US DEPT HEALTH & HUMAN SERVICES.

PI Wang J, Oppenheim JJ, Su S, Gong W, Gao J, Murphy PM;

PI WPI; 2001-211457/21.

PT New serum amyloid A and formyl peptide receptor variant complex and its

PT modulators, useful for treating immune system disorders, amyloidosis,

PT inflammation, infection, organ rejection, arthritis, atherosclerosis

PT and neoplasia -

PS Disclosure; Page 16; 141pp; English.

CC The present sequence is a fragment of human serum amyloid A (SAA). The

CC invention relates to the discovery that SAA is a ligand for FPR1 (human

CC formyl peptide receptor variant). A complex has been isolated that

CC comprises a peptide agent with a sequence corresponding to SAA, or its

CC conservative variant or functional fragment, bound to FPR1. Modulators

CC of the SAA/FPR1 complex are useful for treating immune system

CC disorders, amyloidosis, inflammation, infection, organ rejection,

CC arthritis, atherosclerosis and neoplasia. Fragments of SAA were tested

CC for their ability to modulate assembly of the SAA/FPR1 complex.

SQ Sequence 54 AA;

Query Match 100.0%; Score 23; DB 22; Length 54;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYDA 4  
 ||||  
 Db 42 NYDA 45

## RESULT 31

AAB90295  
 ID AAB90295 standard; Protein; 55 AA.

AC AAB90295;

DT 01-JUN-2001 (first entry)

DE Human SAA carboxy truncation, SEQ ID NO: 53.

KW Human; serum amyloid A; SAA; human formyl peptide receptor variant;

KW FPR1; immunomodulatory; antiinflammatory; antimicrobial; antiarthritic;

KW antiatherosclerotic; immunosuppressive; SAA/FPR1 complex; cancer;

KW immune system disorder; amyloidosis; inflammation; infection;

KW organ rejection; arthritis; atherosclerosis.

OS Homo sapiens.